

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2006, 11:05:32 ; Search time 259 Seconds

(without alignments)  
3418.680 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 MELALCRWGLLALLPPGA.....TFKGTPTAENPEYLGLDVPV 1255

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: uniprot\_05.80:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	6815	100.0	1255	1 ERBB2_HUMAN	P04626 homo sapien
2	6304	92.5	1259	1 ERBB2_CANFA	O18733 canis famil
3	6007.5	88.2	1259	2 O8K3P5_RAT	P06494 rattus norv
4	6003	88.1	1257	1 ERBB2_RAT	P06494 rattus norv
5	5993.5	87.9	1254	1 ERBB2_MESAU	P06053 mesocricetu
6	5982.5	87.8	1305	2 O6ZPE0_MOUSE	O6ZPE0 mus musculu
7	4216	61.9	881	2 O6C0E7_MOUSE	O6C0E7 m mus muscu
8	3514.5	51.6	1275	2 O5EBY4_BRABE	O5EBY4 brachydantio
9	3256.5	47.8	711	2 O80Y89_MOUSE	O80Y89 mus musculu
10	3172	46.5	1209	2 O9GX70_RAT	O9GX70 rattus norv
11	3167	46.5	1210	1 EGFR_HUMAN	P00533 homo sapien
12	3152.5	46.3	1209	2 O8M1L8_PIG	O8M1L8 sus scrofa
13	3146	46.2	1210	1 EGFR_MOUSE	O01279 mus musculu
14	3146	46.2	1210	2 O5SV68_MOUSE	O5SV68 mus musculu
15	3143	46.1	1290	2 O6PPE8_MOUSE	O6PPE8 mus musculu
16	3045	44.7	1292	2 O4PLA5_CHICK	O4PLA5 gallus galli
17	3036	44.5	1276	2 O4PLA4_CHICK	O4PLA4 gallus galli
18	3006	44.1	1302	2 O6UZA8_RAT	O6UZA8 rattus norv
19	3004.5	44.1	1308	1 ERBB4_HUMAN	O15103 homo sapien
20	3004	44.1	1308	2 O6UZA9_RAT	O6UZA9 rattus norv
21	2989	43.9	1308	1 ERBB4_RAT	O62956 rattus norv
22	2988.5	43.9	1091	2 O5O4U8_HUMAN	O5O4U8 homo sapien
23	2986.5	43.8	1191	2 O76ZP7_BRABE	O76ZP7 brachydantio
24	2977.5	43.7	1191	2 O6VQA3_BRABE	O6VQA3 brachydantio
25	2883.5	42.3	1209	2 O6XJY8_XIPXI	O6XJY8 xiphophorus
26	2857	41.9	1081	2 O59FL8_HUMAN	O59FL8 homo sapien
27	2758.5	40.5	1335	2 O4RG29_TETNG	O4RG29 tetraodon n
28	2758	40.5	1165	2 O9YH40_XIPXI	O9YH40 xiphophorus
29	2735.5	40.1	1137	2 O9W6F6_CHICK	O9W6F6 gallus galli
30	2724.5	40.0	1167	1 XM8K_XIPXA	P13388 xiphophorus
31	2620.5	38.5	1094	2 O4RWY4_TETNG	O4RWY4 tetraodon n

32	2441.5	35.8	1342	1 ERBB3_HUMAN	P21860 homo sapien
33	2438.5	35.8	1342	2 O5RB22_PONPY	O5RB22 pongo pygma
34	2408	35.3	1339	2 O6B364_MOUSE	O6B364 mus musculu
35	2370.5	34.8	1339	1 ERBB3_RAT	O62799 rattus norv
36	2326	34.1	1328	2 P79754_FUGRU	P79754 fugu rubrip
37	2282	33.5	1429	2 O5G254_BRABE	O5G254 brachydantio
38	2209.5	32.4	1305	2 O8AW81_BRABE	O8AW81 brachydantio
39	2137.5	31.4	1344	2 O4RW26_TETNG	O4RW26 tetraodon n
40	2102	30.8	892	2 O59EM4_HUMAN	O59EM4 homo sapien
41	2066	30.3	1547	2 O4RXV8_TETNG	O4RXV8 tetraodon n
42	2062.5	30.3	1321	2 O7PPN5_ANOGA	O7PPN5 anopheles g
43	2059.5	30.2	1375	2 O7FHU6_ANOGA	O7FHU6 anopheles g
44	2043.5	30.0	1433	2 O9BIH9_ANOGA	O9BIH9 anopheles g
45	2024.5	29.7	435	2 O6ZMM4_HUMAN	O6ZMM4 homo sapien

#### ALIGNMENTS

RESULT 1  
ERBB2\_HUMAN STANDARD: PRT: 1255 AA.  
ID P04626; Q14256; O6LDV1; Q9UMK4;  
AC 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Receptor tyrosine-protein kinase erbB-2 precursor (BC 2.7.1.112)  
DE (p185erbB2) (c-erbB-2) (NEU proto-oncogene) (Tyrosine kinase-type cell  
DE surface receptor HER2) (MLN 19).  
GN Name=ERBB2; Synonyma=HER2, NEU, NGL;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;  
OC Homo.  
OX NCBI\_Taxid=9606;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=86118663; PubMed=302577;  
RA Yamamoto T., Ikawa S., Akizama T., Samba K., Nomura N., Miyajima N.,  
RA Saito T., Toyoshima K.;  
RT "Similarity of protein encoded by the human c-erbB-2 gene to  
RT epidermal growth factor receptor.";  
RL Nature 319:230-234(1986).  
RN (2)  
RP NUCLEOTIDE SEQUENCE, AND VARIANT ALA-1170.  
RX MEDLINE=86070181; PubMed=2999974;  
RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,  
RA Seeburg P.H., Liberman T.A., Schlessinger J., Francke U.,  
RA Levinson A., Ullrich A.;  
RT "Tyrosine kinase receptor with extensive homology to EGF receptor  
RT shares chromosomal location with neu oncogene.";  
RL Science 230:1132-1139(1985).  
RN (3)  
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS CYS-452; VAL-655 AND  
RP ALA-1170.  
RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,  
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
RA Schickelitz W.S., Sherwood J.K., Wittek L.A., Nickerson D.A.;  
RT "NIBS-SNBs, environmental genome project, NIBS BS1578, Department  
RT of Genome Sciences, Seattle, WA (URL: <http://egp.gs.washington.edu>).";  
RL Submitted (DEC-2002) to the EMBL/Genbank/DBJ databases.  
RN (4)  
RP NUCLEOTIDE SEQUENCE OF 1-191.  
RX MEDLINE=87286898; PubMed=3039351;  
RA Tal M., King C.R., Kraus M.H., Ullrich A., Schlessinger J., Givol D.;  
RT "Human HER2 (neu) promoter: evidence for multiple mechanisms for  
RT transcriptional initiation.";  
RL Mol. Cell. Biol. 7:2597-2601(1987).  
RN (5)  
RP NUCLEOTIDE SEQUENCE OF 737-1031.  
RX MEDLINE=86016729; PubMed=2995967;  
RA Samba K., Kamata N., Toyoshima K., Yamamoto T.;  
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-

RT erbB-1/epidermal growth factor-receptor gene and is amplified in a  
 RT human salivary gland adenocarcinoma".  
 CC Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).  
 CC [6]  
 RN NUCLEOTIDE SEQUENCE OF 832-909.  
 RP TISUE=Mammary carcinoma;  
 RX MEDLINE=85372597; PubMed=2992089;  
 RA King C.R., Kraus M.H., Aaronson S.A.;  
 RT "Amplification of a novel v-erbB-related gene in a human mammary  
 RT carcinoma.";  
 RL Science 229:974-976(1985).  
 RN [7]  
 RN NUCLEOTIDE SEQUENCE OF 1081-1245, AND VARIANT ALA-1170.  
 RP MEDLINE=94000386; PubMed=8104414;  
 RX Sarker F.H., Ball D.B., Li Y.W., Criseman J.D.;  
 RT "Molecular cloning and sequencing of an intron of Her-2/neu (ERBB2)  
 RT gene.";  
 RL DNA Cell Biol. 12:611-615(1993).  
 RN [8]  
 RP IDENTIFICATION IN A COMPLEX WITH PIK3C2A AND EGFR, IDENTIFICATION IN A  
 RP COMPLEX WITH PIK3C2B AND EGFR, AND INTERACTION WITH PIK3C2B.  
 RX PubMed=10805725;  
 RA Arcaro A., Zvelebil M.J., Wallisch C., Ulrich A., Waterfield M.D.,  
 RA Down J.;  
 RT "Class II phosphoinositide 3-kinases are downstream targets of  
 RT activated polypeptide growth factor receptors.";  
 PL Mol. Cell. Biol. 20:3817-3830(2000).  
 RN [9]  
 RP INTERACTION WITH PLXNB1.  
 N.; PubMed=15210733; DOI=10.1083/jcb.200312094;  
 RA Swiercz J.M., Kuner R., Offermanns S.;  
 RT "Plaxin-B1/RhogEF-mediated RhoA activation involves the receptor  
 RT tyrosine kinase ErbB-2.";  
 RL J. Cell Biol. 165:869-880(2004).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 654-662 IN COMPLEX WITH HLA  
 RP AND BETA-2 MICROGLOBULIN.  
 RX MEDLINE=20062861; PubMed=10593938; DOI=10.1074/jbc.274.51.36422;  
 RA Kuhns J., Batalla M.A., Van S., Collins E.J.;  
 RT "Poor binding of a HER-2/neu epitope (GP) to HLA-A2.1 is due to a  
 RT lack of interactions with the center of the peptide.";  
 RL J. Biol. Chem. 274:36422-36427(1999).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.52 ANGSTROMS) OF 23-629 IN COMPLEX WITH FAB.  
 RX MEDLINE=22497871; PubMed=12610629; DOI=10.1038/nature01392;  
 RA Cho H.-S., Mason K., Ramyar K.X., Stanley A.M., Gabelli S.B.,  
 RA Denney D.W., Jr., Leahy D.J.;  
 RT "Structure of the extracellular region of HER2 alone and in complex  
 RT with the Herceptin Fab.";  
 RL Nature 421:756-760(2003).  
 RN [12]  
 RP VARIANTS VAL-654 AND VAL-655.  
 RX MEDLINE=93194196; PubMed=8055488;  
 RA Ehsani A., Low J., Wallace R.B., Wu A.M.;  
 RT "Characterization of a new allele of the human ERBB2 gene by allele-  
 RT specific competition hybridization.";  
 RL Genomics 15:426-429(1993).  
 CC -1- FUNCTION: Essential component of a neurogulin-receptor complex,  
 CC although neurogulins do not interact with it alone. GP30 is a  
 CC potential ligand for this receptor. Not activated by EGF, TGF-  
 CC alpha and amphiregulin.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: Heterodimer with each of the other ERBB receptors  
 CC (potential). Interacts with PRKCAP (By similarity). Binds PLXNB1.  
 CC Part of a complex with EGFR and either PIK3C2A or PIK3C2B. May  
 CC interact with PIK3C2B when phosphorylated on Tyr-1196.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine residues  
 CC (By similarity).  
 CC -1- POLYMORPHISM: There are four alleles due to the variations in  
 CC positions 654 and 655. Allele B1 (Ile-654/Ile-655) has a frequency  
 CC of 0.782; allele B2 (Ile-654/Val-655) has a frequency of 0.206;

CC allele B3 (Val-654/Val-655) has a frequency of 0.012.  
 CC -1- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor  
 CC subfamily.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL, M11767; AAA35808.1; -; Genomic\_DNA.  
 DR EMBL, M11761; AAA35808.1; JOINED; Genomic\_DNA.  
 DR EMBL, M11762; AAA35808.1; JOINED; Genomic\_DNA.  
 DR EMBL, M11763; AAA35808.1; JOINED; Genomic\_DNA.  
 DR EMBL, M11764; AAA35808.1; JOINED; Genomic\_DNA.  
 DR EMBL, M11765; AAA35808.1; JOINED; Genomic\_DNA.  
 DR EMBL, M11766; AAA35808.1; JOINED; Genomic\_DNA.  
 DR EMBL, M11730; AAA75493.1; -; mRNA.  
 DR EMBL, M12036; AAA5978.1; -; Genomic\_DNA.  
 DR EMBL, AY208911; AA018082.1; -; Genomic\_DNA.  
 DR EMBL, X03363; CAA27060.1; -; mRNA.  
 DR EMBL, M16792; AAA58637.1; -; Genomic\_DNA.  
 DR EMBL, M16789; AAA58637.1; JOINED; Genomic\_DNA.  
 DR EMBL, M16790; AAA58637.1; JOINED; Genomic\_DNA.  
 DR EMBL, M16791; AAA58637.1; JOINED; Genomic\_DNA.  
 DR EMBL, L29395; AAA35809.1; -; Genomic\_DNA.  
 DR EMBL, M95667; AAC37531.1; -; Unassigned\_DNA.  
 DR PIR, A24571; A24571.  
 DR PDB, 1N8Z; X-ray; C=23-629.  
 DR PDB, 1OVG; Model; A=737-1031.  
 DR PDB, 1S78; X-ray; A/B=23-646.  
 DR Ensembl; ENSG00000141736; Homo sapiens.  
 DR HGNC; HGNC:3430; ERBB2.  
 DR MIM, 164870; -;  
 DR GO; GO:0016021; C:Integral to membrane; NAS.  
 DR GO; GO:0005886; C:Plasma membrane; NAS.  
 DR GO; GO:0005006; F:Epidermal growth factor receptor activity; NAS.  
 DR GO; GO:0043125; F:ErbB-3 class receptor binding; TAS.  
 DR GO; GO:0046982; F:Protein heterodimerization activity; NAS.  
 DR GO; GO:0004716; F:Receptor signaling protein tyrosine kinase . . . ; TAS.  
 DR GO; GO:0008283; P:Cell proliferation; TAS.  
 DR GO; GO:0007507; P:Heart development; TAS.  
 DR GO; GO:0030879; P:mammary gland development; TAS.  
 DR GO; GO:0007339; P:neurogenesis; TAS.  
 DR GO; GO:0048015; P:phosphoinositide-mediated signaling; NAS.  
 DR GO; GO:0006468; P:Protein amino acid phosphorylation; TAS.  
 DR GO; GO:0045765; P:regulation of angiogenesis; NAS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; NAS.  
 DR InterPro; IPR000494; EGFR\_L.  
 DR InterPro; IPR006211; Furin-like.  
 DR InterPro; IPR006212; Furin repeat.  
 DR InterPro; IPR000719; Prot. Kinase.  
 DR InterPro; IPR001245; Tyr\_Kinase.  
 DR InterPro; IPR008266; Tyr\_Kinase\_AS.  
 DR InterPro; IPR004019; Furin-like; 1.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR Pfam; PF02757; YLP; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PPO00001; Prot\_Kinase; 1.  
 DR SMART; SM00261; FU; 3.  
 DR SMART; SM00219; TYRK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR 3D-structure; ATP-binding; Glycoprotein; Kinase; Multigene family;  
 DR Nucleotide-binding; Phosphorylation; Polymorphism; Receptor; Signal;  
 DR Transferrase; Transmembrane; Tyrosine-protein kinase.  
 DR SIGNAL; 1; 21  
 FT CHAIN 1 1255 Potential.  
 FT TOPO\_DOM 22 652 Extracellular (Potential).  
 FT TRANSMEM 653 675 Potential.  
 FT TOPO\_DOM 676 1255 Cytoplasmic (Potential).

FT	DOMAIN	720	987	Protein kinase.	
FT	NP_BIND	726	734	ATP (By similarity).	
Query Match					
Beat Local Similarity 100.0%; Score 6815; DB 1; Length 1255;					
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MELALACRMGLLALLPP	GAASNOVCTGDMKLR	PASPEYTHLDMRLHYOGGQVVGSL 60	
DB	1	MELALACRMGLLALLPP	GAASNOVCTGDMKLR	PASPEYTHLDMRLHYOGGQVVGSL 60	
QY	61	ELTYLPTNASLSPLODIO	EVQGVYLAHQVAVPLQRL	RIYRGTLQFEDNYALAVLDNG 120	
DB	61	ELTYLPTNASLSPLODIO	EVQGVYLAHQVAVPLQRL	RIYRGTLQFEDNYALAVLDNG 120	
QY	121	DELANNTPTVTGASPGGL	RELQLRLSLTEILKGGVLI	QRNPOLCYODTILMKDIFHKNNQLA 180	
DB	121	DELANNTPTVTGASPGGL	RELQLRLSLTEILKGGVLI	QRNPOLCYODTILMKDIFHKNNQLA 180	
QY	181	LTLIDPNRBRACHPCSP	MGKSGRCMGESSEDCOSL	RTVCAGGCAKCKGRLPTDCHEQC 240	
DB	181	LTLIDPNRBRACHPCSP	MGKSGRCMGESSEDCOSL	RTVCAGGCAKCKGRLPTDCHEQC 240	
QY	241	AACTGPKHSDCLACI	HFHNSGICELHCPALVTYNTD	TPESPNPEGRYTFGASCVTACP 300	
DB	241	AACTGPKHSDCLACI	HFHNSGICELHCPALVTYNTD	TPESPNPEGRYTFGASCVTACP 300	
QY	301	YNYLSTDVSCITLVCP	LHNQEVTAEBGTORCEKCS	APCARVCYGLGMEHLREYRATTSAN 360	
DB	301	YNYLSTDVSCITLVCP	LHNQEVTAEBGTORCEKCS	APCARVCYGLGMEHLREYRATTSAN 360	
QY	361	IOEPFGCKKIFGSLA	FLPSPGDPASNTAPLOP	BOLOYFETLBEITGLYISAMPDSP 420	
DB	361	IOEPFGCKKIFGSLA	FLPSPGDPASNTAPLOP	BOLOYFETLBEITGLYISAMPDSP 420	
QY	421	DLVSFQNLQVIRGR	ILHNGAVSLTLOGLGISWLG	LSRLSELGSLAIHNTHLCEVHTV 480	
DB	421	DLVSFQNLQVIRGR	ILHNGAVSLTLOGLGISWLG	LSRLSELGSLAIHNTHLCEVHTV 480	
QY	481	PMDOFLFRNPHQALL	HTANRPEDECVGEGIA	CARHGICWGPPTQCVCNGSFLRGQEC 540	
DB	481	PMDOFLFRNPHQALL	HTANRPEDECVGEGIA	CARHGICWGPPTQCVCNGSFLRGQEC 540	
QY	541	VEECVLOGLPREYVNA	HHCLPCHPECOPOHNSVTC	CFGEBAOCVCAHYKDPFVCVAC 600	
DB	541	VEECVLOGLPREYVNA	HHCLPCHPECOPOHNSVTC	CFGEBAOCVCAHYKDPFVCVAC 600	
QY	601	PGSVKPDLSYMPIM	WKFPDEEGACQPCPINCT	SHSCVDLDDKGCPEAQASPLTSISAVVG 660	
DB	601	PGSVKPDLSYMPIM	WKFPDEEGACQPCPINCT	SHSCVDLDDKGCPEAQASPLTSISAVVG 660	
QY	661	ILLVVVLGVVFGILL	IKRQOKIRKTYMRLLO	ETELVEPLTPSGAMPNOAKRILKETEL 720	
DB	661	ILLVVVLGVVFGILL	IKRQOKIRKTYMRLLO	ETELVEPLTPSGAMPNOAKRILKETEL 720	
QY	721	RKYKVLGSGAFGVYKGI	WIPIGSENVKIVAIKVL	RENTSPRANKELDEAYVMAGVSP 780	
DB	721	RKYKVLGSGAFGVYKGI	WIPIGSENVKIVAIKVL	RENTSPRANKELDEAYVMAGVSP 780	
QY	781	YYSRLGLICLTSTVOL	VTOLMPYGCGLDHYRENR	GLSGODLLNMCQIAKGSYLEDEV 840	
DB	781	YYSRLGLICLTSTVOL	VTOLMPYGCGLDHYRENR	GLSGODLLNMCQIAKGSYLEDEV 840	
QY	841	LVHRDLAANVILVKS	FNHYKITDPEGLARLLD	IDETEHADGKVPIKMALESILRRRT 900	
DB	841	LVHRDLAANVILVKS	FNHYKITDPEGLARLLD	IDETEHADGKVPIKMALESILRRRT 900	
QY	901	HOSDWSYGVYWEIM	TPGAKPYDGI	PARBIPDLKSGRLP	PPICITDVMIMKCM 960
DB	901	HOSDWSYGVYWEIM	TPGAKPYDGI	PARBIPDLKSGRLP	PPICITDVMIMKCM 960
QY	961	IDSECRPRELIVSE	SFARMADPQRFVYI	QNEDLGASPLDSTFYRSLL	EDDDMGDLVDA 1020
DB	961	IDSECRPRELIVSE	SFARMADPQRFVYI	QNEDLGASPLDSTFYRSLL	EDDDMGDLVDA 1020

Db	961	IDSECRPRELIVSE	SFARMADPQRFVYI	QNEDLGASPLDSTFYRSLL	EDDDMGDLVDA	1020
Qy	1021	EEYLVPOGFCPCD	PAPAGAGVHHRHS	STRSGGDLTTLGLPSEEBEAP	SPPLAPSEB	1080
Db	1021	EEYLVPOGFCPCD	PAPAGAGVHHRHS	STRSGGDLTTLGLPSEEBEAP	SPPLAPSEB	1080
Qy	1081	AGSDVFDGDLGMA	KGLQSLPTHDPSPLQ	RYSEDPVYLPSEETGYVAPLTC	SPQPEYV	1140
Db	1081	AGSDVFDGDLGMA	KGLQSLPTHDPSPLQ	RYSEDPVYLPSEETGYVAPLTC	SPQPEYV	1140
Qy	1141	NOPDVRPOPSPS	REBPLPAARAGATL	ERPKTLSSGKNGVMDVAFGAVEN	PEYLTQ	1200
Db	1141	NOPDVRPOPSPS	REBPLPAARAGATL	ERPKTLSSGKNGVMDVAFGAVEN	PEYLTQ	1200
Qy	1201	GGAPQPPHPPAF	SFADLVYMDOPP	PERGAPSTFKGTPTAENPEYLG	LDVVP	1255
Db	1201	GGAPQPPHPPAF	SFADLVYMDOPP	PERGAPSTFKGTPTAENPEYLG	LDVVP	1255

RESULT 2

ERBB2	CANPA	STANDARD;	PRT;	1259	AA.	
ID	ERBB2	CANPA	STANDARD;	PRT;	1259	AA.
AC	018735;					
DT	25-OCT-2004 (Rel. 45, Created)					
DT	25-OCT-2004 (Rel. 45, Last sequence update)					
DT	13-SEP-2005 (Rel. 48, Last annotation update)					
DE	Receptor tyrosine-protein kinase erbB-2 precursor (EC 2.7.1.112)					
DE	(p155erbB2) (C-erbB-2).					
GN	Name=ERBB2;					
OS	Canis familiaris (Dog).					
OC	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;					
OC	Canis.					
OX	NCBI_TaxID=9615;					
RN	(1)					
RC	NUCLEOTIDE SEQUENCE.					
RC	TISSUE=Mammary gland;					
RA	Yokota H.;					
RT	"cDNA cloning of erbB-2 from canine mammary gland."					
RL	Submitted (OCT-1997) to the EMBL/Genbank/DDBJ databases.					
CC	- FUNCTION: Essential component of a neurotrophin-receptor complex,					
CC	although neurotrophins do not interact with it alone. Gp30 is a					
CC	potential ligand for this receptor. Not activated by EGF, TGF-					
CC	alpha and amphiregulin (By similarity).					
CC	- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein					
CC	tyrosine phosphate.					
CC	- SUBUNIT: Heterodimer with each of the other ERBB receptors					
CC	(potential). Interacts with PRKCAP. Binds PLXNB1. Part of a					
CC	complex with EGFR and either PIK3C2A or PIK3C2B. May interact with					
CC	PIK3C2B when phosphorylated on Tyr-1200 (By similarity).					
CC	- SUBCELLULAR LOCATION: Type I membrane protein.					
CC	- PTM: Ligand-binding increases phosphorylation on tyrosine residues					
CC	(By similarity).					
CC	- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor					
CC	subfamily.					
CC	-----					
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use as long as its content is in no way modified and this statement is not					
CC	removed.					
CC	-----					
CC	EMBL; AB008451; BAA23127.1; -; mRNA.					
CC	HSP; P04626; INB2.					
CC	SMR; O18735; 23-628.					
CC	Ensembl; ENSCAFG0000016351; Canis familiaris.					
CC	InterPro; IPR000494; EGFR_L.					
CC	InterPro; IPR006211; Furin-like.					
CC	InterPro; IPR006212; Furin repeat.					
CC	InterPro; IPR000719; Prot_Kinase.					
CC	InterPro; IPR001245; Tyr_Kinase.					
CC	InterPro; IPR008266; Tyr_Kinase_AS.					
CC	InterPro; IPR004019; YLP_motif.					

DR	Pfam; PF00757; Furin-like; 1.
DR	Pfam; PF01030; Recep_L_domain; 2.
DR	Pfam; PF02757; YLP_2.
DR	PRINTS; PR00109; TYRKINASE.
DR	Prodom; PD000001; Prot_kinase; 1.
DR	SMART; SMO0261; FU; 3.
DR	SMART; SMO0219; TyrcK; 1.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW	ATP-binding; Glycoprotein; Kinase; Multigene family;
KW	Nucleotide-binding; Phosphorylation; Receptor; Signal; Transferase;
KM	Transmembrane; Tyrosine-protein kinase.
FT	SIGNAL 1 22 Potential.
FT	CHAIN 23 1259 Receptor tyrosine-protein kinase erbB-2.
FT	TOPD_DOM 23 653 Extracellular (Potential).
FT	TRANSMEM 654 674 Potential.
FT	TOPD_DOM 675 1259 Cytoplasmic (Potential).
FT	DOMAIN 719 986 Protein kinase.
FT	NP_BIND 725 733 ATP (By similarity).
FT	REGION 1199 1201 Interaction with PIK3C2B (By similarity).
FT	COMPRBS 192 268 Cys-rich.
FT	COMPHAS 1101 1219 Pro-rich.
FT	ACT_SITE 844 844 By similarity.
FT	BINDING 752 752 ATP (By similarity).
FT	MOD_RES 1138 1138 Phosphotyrosine (by autocatalysis) (By similarity).
FT	MOD_RES 1200 1200 Phosphotyrosine (potential).
FT	MOD_RES 1252 1252 Phosphotyrosine (by autocatalysis) (By similarity).
FT	CARBOHYD 68 68 N-linked (GlcNAc . . ) (Potential).
FT	CARBOHYD 259 259 N-linked (GlcNAc . . ) (Potential).
FT	CARBOHYD 421 421 N-linked (GlcNAc . . ) (Potential).
FT	CARBOHYD 529 529 N-linked (GlcNAc . . ) (Potential).
FT	CARBOHYD 570 570 N-linked (GlcNAc . . ) (Potential).
FT	CARBOHYD 628 628 N-linked (GlcNAc . . ) (Potential).
FT	DISULFID 195 204 By similarity.
FT	DISULFID 199 212 By similarity.
FT	DISULFID 220 227 By similarity.
FT	DISULFID 224 235 By similarity.
FT	DISULFID 236 244 By similarity.
FT	DISULFID 240 252 By similarity.
FT	DISULFID 255 264 By similarity.
FT	DISULFID 268 295 By similarity.
FT	DISULFID 299 311 By similarity.
FT	DISULFID 315 331 By similarity.
FT	DISULFID 334 338 By similarity.
FT	DISULFID 511 519 By similarity.
FT	DISULFID 514 527 By similarity.
FT	DISULFID 530 539 By similarity.
FT	DISULFID 543 559 By similarity.
FT	DISULFID 562 575 By similarity.
FT	DISULFID 566 583 By similarity.
FT	DISULFID 586 595 By similarity.
FT	DISULFID 599 622 By similarity.
FT	DISULFID 625 633 By similarity.
FT	DISULFID 629 641 By similarity.
SO	SEQUENCE 1259 AA; 137991 MW; E37364DA9CAACD46 CRC64;
Query Match	
Best Local Similarity 92.5%; Score 6304; DB 1; Length 1259;	
Matches 1162; Conservative 38; Mismatches 54; Indels 6; Gaps 2	
Oy	1 MELAALCRMGILLALLIPGASTOVCTGTDMKLRLPASETHLMDMLRHLYOGCGVQGNL 60
Db	1 MELAAICRWGLLLALLPSGAAGTVCOTGTDMLRLPASETHLMDMLRHLYOGCGVQGNL 60
Oy	61 ELTYVPTNASSFLDDIOGVGVIIAHNVQVPLQRIRIYRGQLPEDNALAVLNG 120
Db	61 ELTYVPANSSLSFLDDIOGVGVIIAHNSQVQPIQRIRIYRGQLPEDNALAVLNG 120
Oy	121 DELANTTPVTGASPGGLRELQLRSLEILIKGVLTIORNPOLCYODTILMKDIFFHNNOLA 180

Db	121	PLRSGFIPAPAAQGGRLQLRSLTEILKGVLIQSSPOLCHODTILMKOVFNKNQLA	180
Qy	181	LTLIDTRSRACCHCSPWCKKSRCWGBSSBDOSLTRYCAGGARCKGRLPTQCHQOC	240
Db	181	LTLIDTRNFSA CPCPSPA CKDAHGWASSGDCSLTRVCAGGARCGPOTDCHQOC	240
Qy	241	AAGCTGPHGSHCLCLAHFNHSGICELCPALVTYNTDTFESMPNDEGRYTTGASCVTACP	300
Db	241	AAGCTGPHGSHCLCLAHFNHSGICELCPALVTYNTDTFESMPNDEGRYTTGASCVTACP	300
Qy	301	YNYISTDVGSCTLYVCP LHNQSVTAE DGTQREKSKPCARVCYLGMEHLEVPATYSAN	360
Db	301	YNYISTDVGSCTLYVCP LHNQSVTAE DGTQREKSKPCARVCYLGMEHLEVPATYSAN	360
Qy	361	IOEFGACKKIGSLAFLPESFDGAPASNTALQEOLOVFETLEETIYLIYSAMPDLP	420
Db	361	IOEFGACKKIGSLAFLPESFDGAPASNTALQEOLOVFETLEETIYLIYSAMPDLP	420
Qy	421	DLASFQMLQVIRGRILHNQAVSLTLQGLGISMGLRSJRELGSGGLALIHNTHLCEVTV	480
Db	421	NLSVFQNLRVIRGRVTLHDGAYSLTLQGLGISMGLRSJRELGSGGLALIHNRARLCFVTV	480
Qy	481	PMDOLFRRPHOALLHTANRPEDECVGBELACHOLCARHCMBGPRPTOCVNSOFLRGQEC	540
Db	481	PMDOLFRRPHOALLHTANRPEDECVGBELACHOLCARHCMBGPRPTOCVNSOFLRGQEC	540
Qy	541	VEBCRVLOGLREYEVYNAHCLPCHPECOPONGSVTFCFPAZDQCVACAHYMDPEFCVARC	600
Db	540	VEBCRVLOGLREYEVYNAHCLPCHPECOPONGSVTFCFPAZDQCVACAHYMDPEFCVARC	599
Qy	601	PSGVKPDLSYMPIMKFPDEBGACOPCPINCTHSCVDLDDKCCPABORASPLTSHSAVVG	660
Db	600	PSGVKPDLSYMPIMKFPDEBGACOPCPINCTHSCVDLDDKCCPABORASPLTSHSAVVG	659
Qy	661	ILLVVLGVNVGIIILKRQOKIRKTYMRRLQETELVEPLPSSGAMPQAOIRLIKETEL	720
Db	660	ILLVVLGVNVGIIILKRQOKIRKTYMRRLQETELVEPLPSSGAMPQAOIRLIKETEL	719
Qy	721	RKVKVLGSGAGTYYKGIWIPDGENVKI PVAIKVIRENTSPKANKEIIDEAYVMAVGSP	780
Db	720	RKVKVLGSGAGTYYKGIWIPDGENVKI PVAIKVIRENTSPKANKEIIDEAYVMAVGSP	779
Qy	781	YVSRLLGICLTSTYQVLTQMLMPYCCCLDHYENENGRGSDOLLMCMQOIAAGMSTLEBYR	840
Db	780	YVSRLLGICLTSTYQVLTQMLMPYCCCLDHYENENGRGSDOLLMCMQOIAAGMSTLEBYR	839
Qy	841	LVHRDLAARNLVYSPNHVKITTDGLARLIDIBETEYADGCKVPIKMMALESIILRRFT	900
Db	840	LVHRDLAARNLVYSPNHVKITTDGLARLIDIBETEYADGCKVPIKMMALESIILRRFT	899
Qy	901	HOSDVMSTGVTVMELMTFGAKPYGIGIPAREIPDLLEKGEKRLPOPPICITDYVMIMVKCM	960
Db	900	HOSDVMSTGVTVMELMTFGAKPYGIGIPAREIPDLLEKGEKRLPOPPICITDYVMIMVKCM	959
Qy	961	IDSECRPFRELVSEPSMARMDPORFVYIQNEIDGCPASPLUSTFASRLIEDDDMDGLVDA	1020
Db	960	IDSECRPFRELVSEPSMARMDPORFVYIQNEIDGCPASPLUSTFASRLIEDDDMDGLVDA	1019
Qy	1021	EEYLVPOQGFCCPDPAAGGVVTHRHRSSTRSGGDDLTGLEPSEBEARSPSLAPSEG	1080
Db	1020	EEYLVPOQGFCCPDPAAGGVVTHRHRSSTRSGGDDLTGLEPSEBEARSPSLAPSEG	1079
Qy	1081	AGSVUFPODLCMGAKGLOSLPTHDPSPLOQVSEDDPVYLPSEDDGYVALPTCSQOPRYV	1140
Db	1080	AGSVUFPODLCMGAKGLOSLPTHDPSPLOQVSEDDPVYLPSEDDGYVALPTCSQOPRYV	1139
Qy	1141	NQPDVPRPSPRSGRPLPAARPACTALER-----PKTILSPKNGVMDVFAFGAVENDE	1195
Db	1140	NQPDVPRPSPRSGRPLPAARPACTALER-----PKTILSPKNGVMDVFAFGAVENDE	1194
Qy	1196	YLTPQGGAAPOPHPPPAFSPAFLDNLTYWDOPPERBGAPESTFKGTPTLAKNBYLGLDLYV	1255
Db	1200	YLAPRGAARPOHPPPAFSPAFLDNLTYWDOPPERBGAPESTFKGTPTLAKNBYLGLDLYV	1255



```
RESULT 3
Q8K3F9 RAT PRELIMINARY; PRT; 1259 AA.
ID Q8K3F9 RAT PRELIMINARY; PRT; 1259 AA.
AC Q8K3F9;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Neu proteinocytocytin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BDIX;
RA Watson P.A., Kim K., Chen K.-S., Gould M.N.;
RL Submitted (May-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY116182; AAM50093.1; -; mRNA.
DR HSSP; P06494; 1N8Y.
DR SMR; Q8K3F9; 26-633.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP_motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 4.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 1259 AA; 139101 MW; B724BDSCC33AE953 CRC64;

Query Match 88.2%; Score 6007.5; DB 2; Length 1259;
Best Local Similarity 88.1%; Pred. No. 28-297;
Matches 1106; Conservative 49; Mismatches 100; Indels 1; Gaps 1;

QY 1 MELIALCGMGLTLLALPGCASTGYCTGDMKRLPASPEPTLMDLRLHYOCQVVGNTL 60
DB 4 MELAWCMGFLALPPEIAGTGYCTGDMKRLPASPEPTLMDLRLHYOCQVVGNTL 63
QY 61 ELTYLPTNASLSFLDIOEVQGYVLIANQVAVPQLRLIVRGTLFEEDNYALAVLDNG 120
DB 64 ELTYVPAVASLSFLDIOEVQGYVLIANQVAVPQLRLIVRGTLFEEDNYALAVLDNR 123
QY 121 DPLANTTPTVT-GASFGGLRELOLRSLITELKGGVLIQRNPOLCYODTILMKDIFKKNQL 179
DB 124 DPQDVVAASTPGRTEGRLRLQLRSLITELKGGVLIQRNPOLCYODMVLAKVFRKNQL 183
QY 180 ALLTLDTNRSRACHPCSPWCKGRGSGESSRDCOSLTRVTSAGAGCARCKGPLPTDCCHQ 239
DB 184 APVDIDTNRSRACPCPCAPACKCKNCKGSEPECCQILITGITCSGCAKCKRGLPTDCCHQ 243
QY 240 CAAGCTGPGHSDCLACLAFNHSIGICELACPALVYNTDTFESMPNPEGRYTFGASCVTAC 299
DB 244 CAAGCTGPGHSDCLACLAFNHSIGICELACPALVYNTDTFESMPNPEGRYTFGASCVTTC 303
```

```
QY 300 PYNLTSTVGSCTLVCPPLAHNOEVTAEADGTQCEKSKPCARVCYGLGMEHLREVAVTSA 359
DB 304 PYNLTSTVGSCTLVCPPLAHNOEVTAEADGTQCEKSKPCARVCYGLGMEHLRGARAITSD 363
QY 360 NIOEPAGCKKIFGSLAFLPESFDGDPASNTVPLQPEOLQVETTELLEITGYLISAMPSTL 419
DB 364 NVOEFGCKKIFGSLAFLPESFDGDPSSGILPLRPEOLQVETTELLEITGYLISAMPSTL 423
QY 420 PDLVSFQNLQVIRGILHNGAVSLTLQGLGSLWGLRLSLRELSGSLALIHNTHLCFVHT 479
DB 424 RDLVSFQNLRLIRGILHNGAVSLTLQGLGSLHSLRLSLRELSGSLALIHNTHLCFVHT 483
QY 480 VPMQDLFNNPRLQALHTANRPEDECEVGLACHQVLCARHGMCPGPTQVNCOSQFLRQOE 539
DB 484 VPMQDLFNNPRLQALHSGNRPEDECEVGLACHQVLCARHGMCPGPTQVNCOSQFLRQOE 543
QY 540 CVEECRVQLGLPREYVNAHCLPCHPECOPONGSVTCGPPADQVACAHYKDPFCVAR 599
DB 544 CVEECRVWKGLPREVSDKRCPLCHPECOPONSBETCFGSSADQACAHYKDSQVAVR 603
QY 600 CPSGVKPDLSYMPIMKPPDEEGACQPCPINCCHSCVDLDERGCPABQPSVTFILATV 659
DB 604 CPSGVKPDLSYMPIMKPPDEEGACQPCPINCCHSCVDLDERGCPABQPSVTFILATV 663
QY 660 GILVAVLVGVFGLIKRQOKIRKYTRRLLOETELVEPLTPSGAMPNOAQMRLKETE 719
DB 664 GVLFLFLIVVVGILIKRRQOKIRKYTRRLLOETELVEPLTPSGAMPNOAQMRLKETE 723
QY 720 LRKVVLTSSGAPGVYKGIWIPDGENVKI PVAIKYLRNTSPKANKELIDRAYVMAGVGS 779
DB 724 LRKVVLTSSGAPGVYKGIWIPDGENVKI PVAIKYLRNTSPKANKELIDRAYVMAGVGS 783
QY 780 PYVSLRLGICLTSTVQVLTOLMPVGCILDHRENRGRIGSODLQWQIAKMSYLEDV 839
DB 784 PYVSLRLGICLTSTVQVLTOLMPVGCILDHRENRGRIGSODLQWQIAKMSYLEDV 843
QY 840 RLVRHDLAARVNLVSPNHVKITDFGLARLLDIDETEHADGKVPYIKMALESILRRRF 899
DB 844 RLVRHDLAARVNLVSPNHVKITDFGLARLLDIDETEHADGKVPYIKMALESILRRRF 903
QY 900 THQSDVMSYGVTWELMTFGAKPYDGI PARBI PDLLERGERLPOPPCTIDVYMTMKW 959
DB 904 THQSDVMSYGVTWELMTFGAKPYDGI PARBI PDLLERGERLPOPPCTIDVYMTMKW 963
QY 960 MIDSECRPRELVESEFSRMAPDPORFVYIQNEDIGPASPILDSTYRSILBEDDMDGLVD 1019
DB 964 MIDSECRPRELVESEFSRMAPDPORFVYIQNEDIGPASPILDSTYRSILBEDDMDGLVD 1023
QY 1020 ABEYLVPOQGFPCPPAPAGAGMVHRRSSSTRSGGDLTLGLEPSEEEAPRSLAPSE 1079
DB 1024 ABEYLVPOQGFPCPPAPAGAGMVHRRSSSTRSGGDLTLGLEPSEEEAPRSLAPSE 1083
QY 1080 GAGSDVPFDGDLQMAKGLQSLPTHDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQBY 1139
DB 1084 GAGSDVPFDGDLQMAKGLQSLPTHDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQBY 1143
QY 1140 VNQDPVRQPPSPRRGCPPLAARPAAGATLEPRITLSPGKNGVAVKVPFAGGAVENREYLT 1199
DB 1144 VNQDPVRQPPSPRRGCPPLAARPAAGATLEPRITLSPGKNGVAVKVPFAGGAVENREYLT 1203
QY 1200 OGGAPOHPHPAPAFADNLVYMDQDPPBRGAPSTFKGTPTAENPEYLGIDVY 1255
DB 1204 REGTASPHHPAPAFADNLVYMDQDPPBRGAPSTFKGTPTAENPEYLGIDVY 1259

RESULT 4
ERBB2 RAT STANDARD; PRT; 1257 AA.
ID ERBB2 RAT STANDARD; PRT; 1257 AA.
AC P06494; Q6P732;
DT 01-JAN-1988 (Rel. 06, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Receptor tyrosine-protein kinase erbb-2 precursor (EC 2.7.1.112)
```

DE (p185erbB2) (C-erbB-2) (NEU proto-oncogene) (Epidermal growth factor  
DE receptor-related protein).  
GN Name=Erbb2; Synonyms=Neu;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RN NUCLEOTIDE SEQUENCE.  
RC TISSUE=Neuroblastoma;  
RX MEDLINE=86118662; PubMed=3945311;  
RA Bargmann C.I., Hung M.-C., Weinberg R.A.;  
RT "The neu oncogene encodes an epidermal growth factor receptor-related  
RT protein."  
RL Nucleur 319:226-230(1986).  
RN [2]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Prostate;  
GN NIH - Mammalian Gene Collection (MGC) project;  
RL Submitted (NOV-2003) to the EMBL/Genbank/DBJ databases.  
RP [3]  
RP NUCLEOTIDE SEQUENCE OF 634-699.  
RX MEDLINE=9203293; PubMed=1682063;  
RA Masui T., Mann A.M., Macatee T.L., Garland E.M., Okamura T.,  
RA Smith R.A., Cohen S.M.;  
RT "Direct DNA sequencing of the rat neu oncogene transmembrane domain  
RT reveals no mutation in urinary bladder carcinomas induced by N-butyl-  
RT N-(4-hydroxybutyl)nitrosamine, N-[4-(5-nitro-2-furyl)-2-  
RT thiazolyl]formamide or N-methyl-N-nitrosourea."  
RL Carcinogenesis 12:1975-1978(1991).  
RN [4]  
RN NUCLEOTIDE SEQUENCE OF 852-905.  
RC TISSUE=Sciatic nerve;  
RX MEDLINE=9122560; PubMed=2025425; DOI=10.1016/0896-6273(91)90167-X;  
RA Lai C., Lemke G.;  
RT "An extended family of protein-tyrosine kinase genes differentially  
RT expressed in the vertebrate nervous system."  
RL Neuron 6:691-704(1991).  
RN [5]  
RN STRUCTURE BY NMR OF 650-668.  
RX MEDLINE=92155181; PubMed=1346763;  
RA Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D.,  
RA Newman R., Crumpton M.J., Sternberg M.J.B., Campbell I.D.;  
RT "Three dimensional structure of the transmembrane region of the proto-  
RT oncogenic and oncogenic forms of the neu protein."  
RL EMBO J. 11:43-48(1992).  
CC -1- FUNCTION: Essential component of a neurogulin-receptor complex,  
CC although neurogulin do not interact with it alone. GP30 is a  
CC potential ligand for this receptor. Not activated by EGF, TGF-  
CC alpha and amphiregulin.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein  
CC tyrosine phosphate.  
CC -1- SUBUNIT: Heterodimer with each of the other ERBB receptors. The  
CC constitutively activated oncogenic variant forms a homodimer.  
CC Interacts with PRKCAP and PLANK. Part of a complex with EGFR and  
CC either PIK3CA or PIK3CB. May interact with PIK3CB when  
CC phosphorylated on Tyr-1198 (By similarity).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine residues  
CC (By similarity).  
CC -1- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor  
CC subfamily.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL; X03362; CA27059.1; ALT\_INIT; mRNA.  
CC DR EMBL; BC061863; AAH61863.1; ALT\_INIT; mRNA.  
CC PIR; A24562; TVRTNU.

DR PDB; 1N8Y; X-ray; C=23-631.  
DR Ensembl; ENSRNOG0000006450; Rattus norvegicus.  
DR RCG; 2561; Erbb2.  
DR GO; GO:0045595; P:regulation of cell differentiation; TAS.  
DR GO; GO:0042127; P:regulation of cell proliferation; TAS.  
DR InterPro; IPR000464; EGRF\_L.  
DR InterPro; IPR006211; Furin-like.  
DR InterPro; IPR006212; Furin\_repeat.  
DR InterPro; IPR000719; Prot\_Kinase.  
DR InterPro; IPR001245; Tyr\_Kinase.  
DR InterPro; IPR008266; Tyr\_Kinase\_AS.  
DR InterPro; IPR004019; YLP\_motif.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF01030; Recep\_L\_domain; 2.  
DR Pfam; PF02757; YLP; 2.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Prot\_kinase; 1.  
DR SMART; SM00261; Fu; 4.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW 3D-structure: ATP-binding; Disease mutation; Glycoprotein; Kinase;  
KW Multigene family; Nucleotide-binding; Phosphorylation; Proto-oncogene;  
KW Receptor; Signal; Transferase; Transmembrane; Tyrosine-protein kinase.  
FT SIGNAL 21  
FT CHAIN 1  
FT TOPO\_DOM 22 1257  
FT TROPO\_DOM 22 654  
FT TRANSMEM 655 677  
FT TOPO\_DOM 678 1257  
FT DOMAIN 722 989  
FT NP\_BIND 728 736  
FT REGION 1197 1199  
FT COMPBIAS 159 369  
FT ACT\_SITE 473 646  
FT ACT\_SITE 847 847  
FT BINDING 755 755  
FT MOD\_RES 1141 1141  
FT MOD\_RES 1198 1198  
FT MOD\_RES 1250 1250  
FT CARBOHYD 68 68  
FT CARBOHYD 188 188  
FT CARBOHYD 260 260  
FT CARBOHYD 532 532  
FT CARBOHYD 573 573  
FT CARBOHYD 631 631  
FT DISULFID 196 205  
FT DISULFID 200 213  
FT DISULFID 221 228  
FT DISULFID 225 236  
FT DISULFID 237 245  
FT DISULFID 241 253  
FT DISULFID 256 265  
FT DISULFID 269 296  
FT DISULFID 300 312  
FT DISULFID 316 332  
FT DISULFID 335 339  
FT DISULFID 513 522  
FT DISULFID 517 530  
FT DISULFID 533 542  
FT DISULFID 546 562  
FT DISULFID 565 578  
FT DISULFID 569 586  
FT DISULFID 589 598  
FT DISULFID 602 625  
FT DISULFID 628 636  
FT DISULFID 632 644  
FT DISULFID 661 661  
FT VARIANT 632 644  
FT CONFLICT 145 145  
FT CONFLICT 505 509  
FT SEQUENCE 1257 AA; 138832 MW; 6129264583011402 CRC64;

Query Match 88.1%; Score 6003; DB 1; Length 1257;  
Best Local Similarity 88.0%; Pred. No. 3,5e-297;  
Matches 1106; Conservative 49; Mismatches 100; Indels 2; Gaps 2;

```

QY 1 MELAALCRWGLLALLPPGAASSTOVCTGDMKRLRLPASPEHLDMLRLHLYGCGQVVGNTL 60
DB 1 MELAAWCRMGFLALLPPGIAGTGVCTGDMKRLRLPASPEHLDMLRLHLYGCGQVVGNTL 60
QY 61 ELTYLPTNAAASLFLDIOGVGVYLIANNQVAVLQRLRYRGQLPEDNALAVLDNG 120
DB 61 ELTYLPTNAAASLFLDIOGVGVYLIANNQVAVLQRLRYRGQLPEDNALAVLDNG 120
QY 121 DPLNNTTAVT-GASPGELRELQRLRLTEILKGVYLIQNNPOLCYDDTLIMKOIFKNNQL 179
DB 121 DPLNNTTAVT-GASPGELRELQRLRLTEILKGVYLIQNNPOLCYDDTLIMKOIFKNNQL 179
QY 121 DPQNVAASTGRTREGRELQRLRLTEILKGVYLIQNNPOLCYDDTLIMKOIFKNNQL 180
DB 121 DPQNVAASTGRTREGRELQRLRLTEILKGVYLIQNNPOLCYDDTLIMKOIFKNNQL 180
QY 180 ALTLIDTRSRACHPSCPMCKGSRGSGSSSEDDQSLRTVCAGGACRCKGRLPTDCHEQ 239
DB 180 ALTLIDTRSRACHPSCPMCKGSRGSGSSSEDDQSLRTVCAGGACRCKGRLPTDCHEQ 239
QY 181 APVDIDTRSRACHPSCPMCKGSRGSGSSSEDDQSLRTVCAGGACRCKGRLPTDCHEQ 240
DB 181 APVDIDTRSRACHPSCPMCKGSRGSGSSSEDDQSLRTVCAGGACRCKGRLPTDCHEQ 240
QY 240 CAAGCTGPKHSDCLALPHNSGICELCPALVYNTPTFESMNPREGRTYFGASCVTAC 299
DB 240 CAAGCTGPKHSDCLALPHNSGICELCPALVYNTPTFESMNPREGRTYFGASCVTAC 299
QY 241 CAAGCTGPKHSDCLALPHNSGICELCPALVYNTPTFESMNPREGRTYFGASCVTAC 300
DB 241 CAAGCTGPKHSDCLALPHNSGICELCPALVYNTPTFESMNPREGRTYFGASCVTAC 300
QY 300 PNYLSTVGSCTVLCPLHNOEVTAEDEGTORCEKSKPCARVCYGLGMEHLREVAATVSA 359
DB 300 PNYLSTVGSCTVLCPLHNOEVTAEDEGTORCEKSKPCARVCYGLGMEHLREVAATVSA 359
QY 301 PNYLSTVGSCTVLCPLHNOEVTAEDEGTORCEKSKPCARVCYGLGMEHLREVAATVSA 360
DB 301 PNYLSTVGSCTVLCPLHNOEVTAEDEGTORCEKSKPCARVCYGLGMEHLREVAATVSA 360
QY 360 NIOEPAGCKKIFGSLAPLPESFGDPASNTAPLOEOLQVETLEITGYLYISAMPDL 419
DB 360 NIOEPAGCKKIFGSLAPLPESFGDPASNTAPLOEOLQVETLEITGYLYISAMPDL 419
QY 361 NVGEDEGCKKIFGSLAPLPESFGDPASNTAPLOEOLQVETLEITGYLYISAMPDL 420
DB 361 NVGEDEGCKKIFGSLAPLPESFGDPASNTAPLOEOLQVETLEITGYLYISAMPDL 420
QY 420 PDLSVFOULQVIRGILHNGAYSLTLOGLISWGLBSLRLREGSLALIHNTLCPHYT 479
DB 420 PDLSVFOULQVIRGILHNGAYSLTLOGLISWGLBSLRLREGSLALIHNTLCPHYT 479
QY 421 RDLSEFOWLRIRGILHNGAYSLTLOGLISWGLBSLRLREGSLALIHNTLCPHYT 480
DB 421 RDLSEFOWLRIRGILHNGAYSLTLOGLISWGLBSLRLREGSLALIHNTLCPHYT 480
QY 480 VPMDOFLPNPQALLHTANRPDE-CVGEGLACHOLCARGHMGWGPCTQVNCOSQFLRG 538
DB 480 VPMDOFLPNPQALLHTANRPDE-CVGEGLACHOLCARGHMGWGPCTQVNCOSQFLRG 538
QY 481 VPMDOFLPNPQALLHTANRPDE-CVGEGLACHOLCARGHMGWGPCTQVNCOSQFLRG 540
DB 481 VPMDOFLPNPQALLHTANRPDE-CVGEGLACHOLCARGHMGWGPCTQVNCOSQFLRG 540
QY 539 ECVEECRVLOGLPREYVNAHQLPCHPECOPOGNSVTGFBADQCVACAHYKDPFCVA 598
DB 539 ECVEECRVLOGLPREYVNAHQLPCHPECOPOGNSVTGFBADQCVACAHYKDPFCVA 598
QY 541 ECVEECRVLOGLPREYVNAHQLPCHPECOPOGNSVTGFBADQCVACAHYKDPFCVA 600
DB 541 ECVEECRVLOGLPREYVNAHQLPCHPECOPOGNSVTGFBADQCVACAHYKDPFCVA 600
QY 599 RCPGSGVKDLSYMPYWKRPDEBGACQPCPINCSTHSCVDLDKGCFAEGRASPLISIAV 658
DB 599 RCPGSGVKDLSYMPYWKRPDEBGACQPCPINCSTHSCVDLDKGCFAEGRASPLISIAV 658
QY 601 RCPGSGVKDLSYMPYWKRPDEBGACQPCPINCSTHSCVDLDKGCFAEGRASPLISIAV 660
DB 601 RCPGSGVKDLSYMPYWKRPDEBGACQPCPINCSTHSCVDLDKGCFAEGRASPLISIAV 660
QY 659 VGILLVVLGVVFGILIKRROOKIRKTYMRRLQETELVEPLTPSGAMPNOAKRIKET 718
DB 659 VGILLVVLGVVFGILIKRROOKIRKTYMRRLQETELVEPLTPSGAMPNOAKRIKET 718
QY 661 VGILLVVLGVVFGILIKRROOKIRKTYMRRLQETELVEPLTPSGAMPNOAKRIKET 720
DB 661 VGILLVVLGVVFGILIKRROOKIRKTYMRRLQETELVEPLTPSGAMPNOAKRIKET 720
QY 719 ELRKVKVLGSGAFGVYKGIWIPDGENYKIPVAILKRENTSPKANKIILDEAVYAVG 778
DB 719 ELRKVKVLGSGAFGVYKGIWIPDGENYKIPVAILKRENTSPKANKIILDEAVYAVG 778
QY 721 ELRKVKVLGSGAFGVYKGIWIPDGENYKIPVAILKRENTSPKANKIILDEAVYAVG 780
DB 721 ELRKVKVLGSGAFGVYKGIWIPDGENYKIPVAILKRENTSPKANKIILDEAVYAVG 780
QY 779 SPYVSRLLIGTCLTSVOLVQTMPLPGCLLDHYREKRGSLGODLLNMCQIAKGSYED 838
DB 779 SPYVSRLLIGTCLTSVOLVQTMPLPGCLLDHYREKRGSLGODLLNMCQIAKGSYED 838
QY 781 SPYVSRLLIGTCLTSVOLVQTMPLPGCLLDHYREKRGSLGODLLNMCQIAKGSYED 840
DB 781 SPYVSRLLIGTCLTSVOLVQTMPLPGCLLDHYREKRGSLGODLLNMCQIAKGSYED 840
QY 839 VRLVHRDLAARNVLYKSPNHVKTIDPGLARLLIDETRYHADGGKVPKIMMALISILRR 898
DB 839 VRLVHRDLAARNVLYKSPNHVKTIDPGLARLLIDETRYHADGGKVPKIMMALISILRR 898
QY 841 VRLVHRDLAARNVLYKSPNHVKTIDPGLARLLIDETRYHADGGKVPKIMMALISILRR 900
DB 841 VRLVHRDLAARNVLYKSPNHVKTIDPGLARLLIDETRYHADGGKVPKIMMALISILRR 900
QY 899 FTHOSDVMSYGVTWELMTFGAKPYDGI-PAREIPDLLEKGBRLPOPPICTIDVYIMVYC 958
DB 899 FTHOSDVMSYGVTWELMTFGAKPYDGI-PAREIPDLLEKGBRLPOPPICTIDVYIMVYC 958
QY 901 FTHOSDVMSYGVTWELMTFGAKPYDGI-PAREIPDLLEKGBRLPOPPICTIDVYIMVYC 960
DB 901 FTHOSDVMSYGVTWELMTFGAKPYDGI-PAREIPDLLEKGBRLPOPPICTIDVYIMVYC 960
QY 959 WMIDSECPRRRELVSERKARADQQRVYVIONEDLGASPLDSTFFYSLLDEDDMGVLV 1018
DB 959 WMIDSECPRRRELVSERKARADQQRVYVIONEDLGASPLDSTFFYSLLDEDDMGVLV 1018
QY 961 WMIDSECPRRRELVSERKARADQQRVYVIONEDLGASPLDSTFFYSLLDEDDMGVLV 1020
DB 961 WMIDSECPRRRELVSERKARADQQRVYVIONEDLGASPLDSTFFYSLLDEDDMGVLV 1020

```

```

QY 1019 DABEYLVOQGFPCPDPAAGVWVHRNRSSSTRSGGDLITGLPSEEBEAPSLAPS 1078
DB 1021 DABEYLVOQGFPCPDPAAGVWVHRNRSSSTRSGGDLITGLPSEEBEAPSLAPS 1080
QY 1079 EGAGSDVVDGDLGKAGAGLQSLPTHDPSPLOYSSEDTVPLPSTDTGVAPLTSPOPE 1138
DB 1081 EGAGSDVVDGDLGKAGAGLQSLPTHDPSPLOYSSEDTVPLPSTDTGVAPLTSPOPE 1140
QY 1139 VYNOPDVARPOPSPREGLPAPAPAGATLERPKTSLPKGNQVWVDFAPGAVENPEYLT 1198
DB 1141 VYNOPDVARPOPSPREGLPAPAPAGATLERPKTSLPKGNQVWVDFAPGAVENPEYLT 1200
QY 1199 POGGAPOPHPPAPSPAFDNLVYWDQDPPEKAPPTFKGPTAENPEYLGLDVVP 1255
DB 1201 PREGTAPPHPSPAFSPAFDNLVYWDQDSSQGPSPSFEGETPAENPEYLGLDVVP 1257

RESULT 5
ERBB2_MESAU STANDARD; PRT; 1254 AA.
AC 060553;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Receptor tyrosine-protein kinase erbB-2 precursor (BC 2.7.1.112)
DE (p185erbB2) (C-erbB-2) (NEU proto-oncogene).
GN Name=ERBB2; Synonyms=NEU;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Cricetidae; Cricetinae; Mesocricetus.
OX NCBI_TaxId=10036;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Nerve;
RX MEDLINE=94193007; PubMed=7908275; DOI=10.1016/0378-1119(94)90553-3;
RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M., Yamazaki Y.,
RA Ishikawa T.;
RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";
RL Gene 140:251-255(1994).
CC -! FUNCTION: Essential component of a neurotrophin-receptor complex,
CC although neurotrophins do not interact with it alone. Grp30 is a
CC potential ligand for this receptor. Not activated by EGF, TGF-
CC alpha and amphiregulin (By similarity).
CC -! CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein
CC tyrosine phosphate.
CC -! SUBUNIT: Heterodimer with each of the other ERBB receptors
CC (Potential1). Part of a complex with EGFR and either PIK3C2A or
CC PIK3C2B. Interacts with PRKCA and PIKXBL. May interact with
CC PIK3C2B when phosphorylated on Tyr-1196 (By similarity).
CC -! SUBCELLULAR LOCATION: Type I membrane protein.
CC -! PTM: Ligand-binding increases phosphorylation on tyrosine
CC residues.
CC -! SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor
CC subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL: D16295; BAA03801.1; -, mRNA.
CC PIR: I48161; I48161.
CC
CC HSSP: P06494; INBY.
CC
CC SMR: Q60553; 23-629.
CC
CC InterPro: IPR000494; EGFR_L.
CC InterPro: IPR006211; Furin-like.
CC InterPro: IPR006212; Furin repeat.
CC InterPro: IPR000719; Prot_Kinase.
CC InterPro: IPR001245; Tyr_Kinase.
CC InterPro: IPR008266; Tyr_kinase_AS.
CC
CC InterPro: IPR004019; YLP_motif.

```

DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF01030; Recep\_L\_domain; 2.  
DR Pfam; PF02757; YLP; 2.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00261; FU; 4.  
DR SMART; SM00219; TyrcK; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
KW ATP-binding; Disease mutation; Glycoprotein; Kinase; Multigene family;  
KW Nucleotide-binding; Phosphorylation; Proto-oncogene; Receptor; Signal;  
KW Transferrase; Transmembrane; Tyrosine-protein kinase.  
FT SIGNAL 1 21 Potential.  
FT CHAIN 22 1254 Receptor tyrosine-protein kinase erbB-2.  
FT TOPO\_DOM 22 652 Extracellular (Potential).  
FT TRANSDOM 653 675 Potential.  
FT TOPO\_DOM 676 1254 Cytoplasmic (Potential).  
FT DOMAIN 720 987 Protein kinase.  
FT NP\_BIND 726 734 ATP (By similarity).  
FT REGION 1195 1197 Interaction with PIK3C2B (By similarity).  
FT COMPBIAS 158 368 Cys-rich.  
FT ACT\_SITE 845 845 By similarity.  
FT BINDING 753 753 ATP (By similarity).  
FT MOD\_RES 1139 1139 Phosphotyrosine (by autocatalysis) (By similarity).  
FT MOD\_RES 1139 1139 Phosphotyrosine (Potential).  
FT MOD\_RES 1247 1247 Phosphotyrosine (by autocatalysis) (By similarity).  
FT CARBOHYD 68 68 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 125 125 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 187 187 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 259 259 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 530 530 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 571 571 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 629 629 N-linked (GlcNAc...) (Potential).  
FT DISULFID 195 204 By similarity.  
FT DISULFID 199 212 By similarity.  
FT DISULFID 236 244 By similarity.  
FT DISULFID 240 252 By similarity.  
FT DISULFID 255 264 By similarity.  
FT DISULFID 268 295 By similarity.  
FT DISULFID 299 311 By similarity.  
FT DISULFID 315 331 By similarity.  
FT DISULFID 334 338 By similarity.  
FT DISULFID 511 520 By similarity.  
FT DISULFID 515 528 By similarity.  
FT DISULFID 531 540 By similarity.  
FT DISULFID 544 560 By similarity.  
FT DISULFID 563 576 By similarity.  
FT DISULFID 567 584 By similarity.  
FT DISULFID 587 596 By similarity.  
FT DISULFID 600 623 By similarity.  
FT DISULFID 626 634 By similarity.  
FT DISULFID 630 642 By similarity.  
FT VARIANT 658 658 V -> E (in oncogenic NEU).  
FT VARIANT 659 659 V -> E (in oncogenic NEU).  
SQ SEQUENCE 1254 AA; 138253 MW; 974C3791C21F2BE1 CRC64;

Query Match 87.9%; Score 5993.5; DB 1; Length 1254;  
Best local Similarity 87.7%; Pred. No. 1,1e-296;  
Matches 1101; Conservative 57; Mismatches 96; Indels 1; Gaps 1;

QY 1 METALALGKGLLALPRGASTQVCTGDMKRLPASPEHTLDMLEHLYGCGVQGNL 60  
DB 1 METLAAGMGWGLLALLSPGASGTQVCTGDMKRLPASPEHTLDIYVHLLYQGCGVQGNL 60  
QY 61 ELATYPTNASTLSFTODIOEVGVYLIAHQVYPLQRLIRIVRGSTOLFEDNYALAVDNG 120  
DB 61 ELATYLPNASTLSFTODIOEVGVYLIAHQVYPLQRLIRIVRGSTOLFEDNYALAVDNR 120  
QY 121 DPLNNTTAVTGASPGALREIQRLSTELIKGGLIQRNPOLCYDDTILMKDIFHKNNOLA 180

DB 121 DPLNNTTAVTGARTEGALREIQRLSTELIKGGLIQRNPOLCYDDTILMKDIFHKNNOLA 180  
QY 181 LTLIDTNRBACPCSPCKGSRCKMSESSQCSLRTVCGAGCARCGPLPTCCCHQC 240  
DB 181 PVIDITNRBACPCACACCKDNCKGASPEDCQLTGTIAPRAVPAARALPTCCCHQC 240  
QY 241 AAGCTGPKHSDCLALHPNHSIGICELHCPALVTYNTDTFBSMPNDEGRYTFGASCVTACP 300  
DB 241 AAGCTGPKHSDCLALHPNHSIGICELHCPALVTYNTDTFBSMPNDEGRYTFGASCVTACP 300  
QY 301 YNTISTVSGCTLVCPILHNGEYTAEDTORCEKSKCARCYGLGMEHLREVAAYTSAN 360  
DB 301 YNTISTVSGCTLVCPILHNGEYTAEDTORCEKSKCARCYGLGMEHLREVAAYTSAN 360  
QY 361 IOEPAGCKKIFGSLAFIPESFDGPAANTAPLOEQOVFETLEIGYLYISAMPDLP 420  
DB 361 IOEPAGCKKIFGSLAFIPESFDGPAANTAPLOEQOVFETLEIGYLYISAMPDLP 420  
QY 421 DLSVFQNLAVYIRGRILHNGAYSLTLQGLGI SWLGRLSRLBSGLALIHNTHLCPVHTV 480  
DB 421 DLSVFQNLAVYIRGRILHNGAYSLTLQGLGI SWLGRLSRLBSGLALIHNTHLCPVHTV 480  
QY 481 PMDQLFNNPQALLHTANREDECEVSGILACHQLCARGHCMPGPTQCVCNCSOTLRQEC 540  
DB 481 PMDQLFNNPQALLHTANREDECEVSGILACHQLCARGHCMPGPTQCVCNCSOTLRQEC 540  
QY 541 VEECRVLOGLPREVYVNRHCLPCHPECPONGSVTCGFEADQCVACAHYKDPFCVARC 600  
DB 541 VEECRVLOGLPREVYVNRHCLPCHPECPONGSVTCGFEADQCVACAHYKDPFCVARC 600  
QY 601 PSQVKPDLSTYMPIMKPFDEBGAQCPDINCTHSCVDLDKGCAPAEQASPLTISIAYVG 660  
DB 601 PSQVKPDLSTYMPIMKPFDEBGAQCPDINCTHSCVDLDKGCAPAEQASPLTISIAYVG 660  
QY 661 ILVYVTVGVVFGILIKRROKIRKYMTRLLQETELVEPLTPSGAMNQOMRLTKETEL 720  
DB 661 ILVYVTVGVVFGILIKRROKIRKYMTRLLQETELVEPLTPSGAMNQOMRLTKETEL 720  
QY 721 RKYVLSGAGFGYVYKGIWIPDGENYKI PVALIKVLRNTPSKANKELIDEAYVYAGVSP 780  
DB 721 RKYVLSGAGFGYVYKGIWIPDGENYKI PVALIKVLRNTPSKANKELIDEAYVYAGVSP 780  
QY 781 YVSLRLGICLTSTVQLVLTQMPYGCILLDHVENRGRIGSODLNMCKQIAKMSYLEYVR 840  
DB 781 YVSLRLGICLTSTVQLVLTQMPYGCILLDHVENRGRIGSODLNMCKQIAKMSYLEYVR 840  
QY 841 LVHRDLAARVVLVKSPPHVKITDGLARLLDIBETEVHADGKVPIMWIALESILRRFT 900  
DB 841 LVHRDLAARVVLVKSPPHVKITDGLARLLDIBETEVHADGKVPIMWIALESILRRFT 900  
QY 901 HQSDVMSYGVTVMEIMTFGAKPYDGI PAREIPDLLEKGERLPOPPICITIDVYIMVCMW 960  
DB 901 HQSDVMSYGVTVMEIMTFGAKPYDGI PAREIPDLLEKGERLPOPPICITIDVYIMVCMW 960  
QY 961 IDSECRPRFRELVSFBSRMARDPQRFVVIQNEIDGPASPLDSTFYRSLDEDDMGDLVDA 1020  
DB 961 IDSECRPRFRELVSFBSRMARDPQRFVVIQNEIDGPASPLDSTFYRSLDEDDMGDLVDA 1020  
QY 1021 EETLYVPQGFPPDPAPGASTAHRRRSSTBSGGELTLGMEPSGEBPRSPSLASEG 1080  
DB 1021 EETLYVPQGFPPDPAPGASTAHRRRSSTBSGGELTLGMEPSGEBPRSPSLASEG 1080  
QY 1081 AGSDVPFGDGLMGAKLQSLPTHPDSPLRYSBSPVPLPSETIDGVYAPLTCSPQREYV 1140  
DB 1081 AGSDVPFGDGLMGAKLQSLPTHPDSPLRYSBSPVPLPSETIDGVYAPLTCSPQREYV 1140  
QY 1141 NQPDVBPQPSPREGPLPAARPGATLERPKTSLPGKNGVYKQVAFGAVENDEYLPQ 1200  
DB 1141 NQPDVBPQPSPREGPLPAARPGATLERPKTSLPGKNGVYKQVAFGAVENDEYLPQ 1200  
QY 1201 GGAAPQHPHPPASPAVDNLTYYWDQPPERGAAPSTTKGPTLANPFLGLDQV 1255  
DB 1201 GGAAPQHPHPPASPAVDNLTYYWDQPPERGAAPSTTKGPTLANPFLGLDQV 1255

Db 1201 GGSASQPH-PPALCPADNLYWQDPPSEBGSPPPTFEGTPTABENPEYLGLDVPV 1254

RESULT 6

Q62PE0\_MOUSE PRELIMINARY; PRT; 1305 AA.

AC Q62PE0; 05-JUL-2004 (Tremblrel. 27, Created)

DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)

DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)

DE MKIAA3023 protein (Fragment).

GN Name=Erbb2; Synonym=mkIAA3023;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muridae; Muridae; Murinae; Mus.

OX NCBI\_TaxId=10090;

RA NUCLEOTIDE SEQUENCE.

RP TISSUE=Embryonic tail;

RX MEDLINE=22977043; PubMed=14621295;

RA Okazaki N., Kikuno R., Ohara K., Inamoto S., Koseki H., Hirooka S., Saga Y., Nagase T., Ohara O., Koga H.;

RT "Prediction of the coding sequences of mouse homologues of KIAA gene: RT i1. the complete nucleotide sequences of 500 mouse KIAA-homologous RT cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries.";

RL DNA Ref. 10:167-180(2003).

DR EMBL; AK129487; BAC98297.1; -; mRNA.

DR HSSP; P13362; 1AGM.

DR SNR; Q62PE0; 72-679.

DR Ensemble; ENSMUSG0000062312; Mus musculus.

DR MGI; MGI:95410; Erbb2.

DR GO; GO:0016324; C:apical plasma membrane; IDA.

DR GO; GO:0005737; C:cyclopia; IDA.

DR GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0007507; P:heart development; IMP.

DR GO; GO:0042552; P:myelination; IMP.

DR GO; GO:0007422; P:peripheral nervous system development; IMP.

DR InterPro; IPR000494; EGFR\_L.

DR InterPro; IPR006211; Furin-like.

DR InterPro; IPR006212; Furin repeat.

DR InterPro; IPR000719; Prot. Kinase.

DR InterPro; IPR002290; Ser. Thr. Kinase.

DR InterPro; IPR001245; Tyr. Kinase.

DR InterPro; IPR008266; Tyr. Kinase\_AS.

DR InterPro; IPR004019; YLP motif.

DR Pfam; PF00757; Furin-like; 1.

DR Pfam; PF01030; Recep\_L\_domain; 2.

DR Pfam; PF02757; YLP\_3.

DR PRINTS; PR00109; TYRKINASE.

DR PRODOM; PD000001; Prot\_Kinase; 1.

DR SMART; SMO0261; FU; 4.

DR SMART; SMO0220; S\_TKc; 1.

DR SMART; SMO0219; Tyrc; 1.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.

DR ATP-binding; Kinase; Nucleotide-binding; Transferase.

FT NON TER 1 1

SEQUENCE 1305 AA; 143508 MW; A5ID897408521860 CRC64;

Query Match 87.8%; Score 5982.5; DB 2; Length 1305;

Best Local Similarity 87.7%; Pred. No. 4e-296;

Matches 1101; Conservative 55; Mismatches 99; Indels 1; Gaps 1;

Qy 1 MELALACMGWGLLLLPAGASTQVCTGTDKMLRLPASPETHLMDLRLYGGCCVVGQNL 60

Db 50 MELALACMGWGLLLLPAGASTQVCTGTDKMLRLPASPETHLMDLRLYGGCCVVGQNL 109

Qy 61 ELTYLPTNASLSTFDIOGVGVYLIANNOVROVPLQRLIRVGTQLFEDNYALAVLDNG 120

Db 110 ELTYLPTNASLSTFDIOGVGVYLIANNOVROVPLQRLIRVGTQLFEDNYALAVLDNR 169

Qy 121 DPLANN-TPYVAGSPGGLREQLRLSLTEILKGVLIQRNPOLCYODTILMKDIFKNNQL 179

Db 170 DPLDNTTAAAGRRTPBEGRLQLRLSLTEILKGVLIQRNPOLCYODTILMKDIFKNNQL 229

Qy 180 ALTLIDNRSAPACHPSPMCKSGSRGSSSDQSLRTVACAGCARCKGLPTDCHEQ 239

Db 220 APVMDITVRSRACPPCAPTCYKDNHCWGBSPBDQTLTGTICTSGCARCKGLPTDCHEQ 289

Qy 240 CAAGCTGPKSDCLACLFHNSGICELHCPALVYNTDTPFSMPNBERYFGASCYAC 299

Db 290 CAAGCTGPKSDCLACLFHNSGICELHCPALVYNTDTPFSMPNBERYFGASCYAC 349

Qy 300 PNYLSTVGSCTVLCPLANOBYTAEDTGRCEKSKPCARVCYGLGMEHLREVAVTS 359

Db 350 PNYLSTVGSCTVLCPLANOBYTAEDTGRCEKSKPCARVCYGLGMEHLREVAVTS 409

Qy 360 NIOEPAGCKKI FGSIAFLPESFDGDPASNTALPQEOLOVEFTLEITGYLYISAMPDL 419

Db 410 NIOEPAGCKKI FGSIAFLPESFDGDPASNTALPQEOLOVEFTLEITGYLYISAMPDL 469

Qy 420 PDLSTVQNLQYIRGRILNNGAVSLTQGLQISWGLRSLRELGGALIHNTILCFYHT 479

Db 470 QDLSTVQNLQYIRGRILNNGAVSLTQGLQISWGLRSLRELGGALIHNTILCFYHT 529

Qy 480 VPMQDLFNPQHALLHTANRPEDCEVGEGLACHOGLCARGHGCMGPGPTCCVNCOSPLRGQ 539

Db 530 VPMQDLFNPQHALLHTANRPEDCEVGEGLACHOGLCARGHGCMGPGPTCCVNCOSPLRGQ 589

Qy 540 CVBECRVLOGLPREYVNAHRLCPHPECOPNGSVTCGPBPADQVACAHYKDPFCYAR 599

Db 590 CVBECRVLOGLPREYVNAHRLCPHPECOPNGSVTCGPBPADQVACAHYKDPFCYAR 649

Qy 600 CPSCGVKPLSLMPYWKPPDEEGACQPCPICTHSCVLDLDDGCPAEOQASPLTSLAVV 659

Db 650 CPSCGVKPLSLMPYWKPPDEEGACQPCPICTHSCVLDLDDGCPAEOQASPLTSLAVV 709

Qy 660 GILLVVLVGVVFGILIKRQOKIRRYTRRLLOETELVEPILTPSGAMPNQOMRLKETE 719

Db 710 GILLVVLVGVVFGILIKRQOKIRRYTRRLLOETELVEPILTPSGAMPNQOMRLKETE 769

Qy 720 LRKVKVLSGAGFVYVYKGIWIPDGENYKIPALIKVLRNTSPKANKELIDEAYVAVGVS 779

Db 770 LRKVKVLSGAGFVYVYKGIWIPDGENYKIPALIKVLRNTSPKANKELIDEAYVAVGVS 829

Qy 780 PYVSRLLGICLTSTVQVLTQMPYGCILDHTRKRGRLGSDLLNMCQIAKGSYLEDV 839

Db 830 PYVSRLLGICLTSTVQVLTQMPYGCILDHTRKRGRLGSDLLNMCQIAKGSYLEDV 889

Qy 840 RLVRHDLAARVLYKSPNHVKITDGLARLIDIDETRYHADGGKVPYKMMALSTLRBRF 899

Db 890 RLVRHDLAARVLYKSPNHVKITDGLARLIDIDETRYHADGGKVPYKMMALSTLRBRF 949

Qy 900 THQSDVMSYGVTVLMTFGAKPYDGIIPARIPDLLEKGERLPORPCTITDVYIMVNCW 959

Db 950 THQSDVMSYGVTVLMTFGAKPYDGIIPARIPDLLEKGERLPORPCTITDVYIMVNCW 1009

Qy 960 MIDSECRPRFRELVSERSMARDPQRFVYVJNEDLGPASPLDSTFTRSLLEDDMDGLVD 1019

Db 1010 MIDSECRPRFRELVSERSMARDPQRFVYVJNEDLGPASPLDSTFTRSLLEDDMDGLVD 1069

Qy 1020 AAEYLVPOQGFPCDDPAAGAGMWHHRHRS6STSGGSDLTGLGPESEEARSLAPSE 1079

Db 1070 AAEYLVPOQGFPCDDPAAGAGMWHHRHRS6STSGGSDLTGLGPESEEARSLAPSE 1129

Qy 1080 GAGSDVFDGDLGMAAGKAGLQSLPTHDSPLQRYSEDPVPLPSESTDGVAVLPTCSPODEY 1139

Db 1130 GAGSDVFDGDLGMAAGKAGLQSLPTHDSPLQRYSEDPVPLPSESTDGVAVLPTCSPODEY 1189

Qy 1140 VNQEDVRRPQPSPPBEGPLPAARPGATLBERKXTLSPGKNGYVYKQVAFAGAVENPEYITP 1199

Db 1190 VNQEDVRRPQPSPPBEGPLPAARPGATLBERKXTLSPGKNGYVYKQVAFAGAVENPEYITP 1249

QY 1200 OGGAAPHPHPAPSPAFDNLVYWDQPPERKAPSTFKGTPTAENBEYLGLDVPV 1255  
Db 1250 RAGTASGPHSPAPSPAFDNLVYWDQSSSEGGPPSTFEGTPTAENBEYLGLDVPV 1305  
RESULT 7  
ID Q8COE7 MOUSE PRELIMINARY; PRT; 881 AA.  
AC Q8COE7;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Mus musculus 13 days embryo male testis cDNA, RIKEN full-length  
DE enriched library, clone:603049F08 product:v-erb-b2 erythroblastic  
DE leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene  
DE homolog (avian), full insert sequence. (Fragment).  
GN Name=ErbB2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Mech. Enzymol. 303:19-44(1999).  
[2]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
RA Knehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,  
RA Schirni L.M., Scahill F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barin G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Ruedenrich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,  
RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
[3]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RA The PANTOM Consortium,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
[4]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,  
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
[5]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
[6]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozawa T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitt (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK031542; BAC27442.1; -; mRNA.  
DR HSSR: P06494; INBY.  
DR SMR: Q8COE7; 1-255.  
DR Ensembl: ENSMUSG00000062312; Mus musculus.  
DR MGI: 95410; ErbB2.  
DR GO: GO:0016324; C:apical plasma membrane; IDA.  
DR GO: GO:0005737; C:cytosol; IDA.  
DR GO: GO:0005515; F:protein binding; IPI.  
DR GO: GO:0007507; P:heart development; IMP.  
DR GO: GO:0042552; P:myelination; IMP.  
DR GO: GO:0007422; P:peripheral nervous system development; IMP.  
DR InterPro: IPR000494; EGFR\_L.  
DR InterPro: IPR006212; Furin\_repeat.  
DR InterPro: IPR000719; Prot\_kinase.  
DR InterPro: IPR001245; Tyr\_kinase.  
DR InterPro: IPR008266; Tyr\_kinase\_AS.  
DR Pfam: PF01030; YLP\_motif.  
DR Pfam: PF02757; YLP\_2.  
DR PRINTS: PR00109; TYRKINASE.  
DR ProDom: PD000001; Prot\_kinase; 1.  
DR SMART: SM00261; FU\_2.  
DR SMART: SM00219; Tyrc; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW ATP-binding; Kinase; Nucleotide-binding; Transferase;  
KW Tyrosine-protein kinase.  
FT NON\_TER 1  
SQ SEQUENCE 881 AA; 97501 MW; 5D5042BE9F80836 CRC64;  
Query Match 61.9%; Score 4216; DB 2; Length 881;  
Best Local Similarity 88.3%; Pred. No. 2 2e-206;  
Matches 778; Conservative 39; Mismatches 64; Indels 0; Gaps 0;  
QY 375 AFLPESFDGDPASNTAPLOEOLQVEFTLEITGYLYISAMPDLSPLDLSYFONLQVIRGR 434  
Db 1 AFLPESFDGDPASNTAPLOEOLQVEFTLEITGYLYISAMPDLSPLDLSYFONLQVIRGR 60  
QY 435 ILNNGAYSLTLOGLGSLWGLRLRLSGSLALIHNTHTCFPHYTPWMDLPNNQALL 494  
Db 61 ILNNGAYSLTLOGLGSLWGLRLRLSGSLALIHNTHTCFPHYTPWMDLPNNQALL 120  
QY 495 HTANRPEDECVGSLACHOLCARGHCWGPPTQVNCSCQPLRGOECVBEGRVAGLPREY 554  
Db 121 HSNRPEDECVGSLACHOLCARGHCWGPPTQVNCSCQPLRGOECVBEGRVAGLPREY 180

```

QY 555 VNARHCLPCHPECOPNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMTW 614
D 181 VRGRKCLPCHPECOPNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMTW 240
QY 615 KPDEEGACQPCPINCTHSVDLDDKCPAEGORASPLTSTISAVVGLLVVAVGFGIL 674
D 241 KYPDEEGICQPCPINCTHSVDLDDKCPAEGORASPLTSTISAVVGLLVVAVGFGIL 300
QY 675 IKRROKIRKYMRLLOETELVEPLTSGAMPNOMRLKETELRKYKVLGSGAFGV 734
D 301 IKRRROKIRKYMRLLOETELVEPLTSGAMPNOMRLKETELRKYKVLGSGAFGV 360
QY 735 YKGIWIPGENVKIPVAIKVLENTSPKANKELIDEAVYMGVSPYVSRLLGLCTSTV 794
D 361 YKGIWIPGENVKIPVAIKVLENTSPKANKELIDEAVYMGVSPYVSRLLGLCTSTV 420
QY 795 QLVTLMPYGLLDHVENRGRSLSQDILNMCQIAKGMSTYEDVLYHRDLAANVLVK 854
D 421 QLVTLMPYGLLDHVENRGRSLSQDILNMCQIAKGMSTYEDVLYHRDLAANVLVK 480
QY 855 SPNHVKINDPGLARLDIDETEHYADGGKVPKGMALSTILRRRTHOSDVMSYGVYWE 914
D 481 SPNHVKINDPGLARLDIDETEHYADGGKVPKGMALSTILRRRTHOSDVMSYGVYWE 540
QY 915 LMTFGAKPYDGIPIAREIPDLLEKGRRLPOPPICTIDVYMIWCKMMIDSECRPRELVS 974
D 541 LMTFGAKPYDGIPIAREIPDLLEKGRRLPOPPICTIDVYMIWCKMMIDSECRPRELVS 600
QY 975 EFSRMARDPQRFVVIQNEEDLGPASPLDSTFYRSLLLEDDMDGLVDAEYLVPQGFPCPD 1034
D 601 EFSRMARDPQRFVVIQNEEDLGPASPLDSTFYRSLLLEDDMDGLVDAEYLVPQGFPCPD 660
QY 1035 PAPAGGVNHHRRSSSTRSGGDLTLGLEPSEERAPSPILASRGASDVVDGDLGMA 1094
D 661 PALGGSFAHRRHRSASRGGGLTLGLEPSEERAPSPILASRGASDVVDGDLGMA 720
QY 1095 AKGLQSLPTHPSPLORYSEDPVPLPSETDGYVAPLTCSPQPEVYNQDVAPQPPSPRE 1154
D 721 TKGLOSLPPLHSLPQRYSEDPVPLPSETDGYVAPLTCSPQPEVYNQDVAPQPPSPRE 780
QY 1155 GPLPARAGATLERRKTLSPGKGVVAVDFAFGAVENPEYLTPQGAAPQHPPPAFS 1214
D 781 GPLPARAGATLERRKTLSPGKGVVAVDFAFGAVENPEYLTPQGAAPQHPPPAFS 840
QY 1215 PAFDNLVYWDQPPPERGAPSTFKCTPTANPEYLGDLV 1255
D 841 PAFDNLVYWDQPPPERGAPSTFKCTPTANPEYLGDLV 881

```

```

RA Pogoda H.-M., Voas M.G., Talbot W.S.;
RT "Analysis of a zebrafish erb2 gene";
D Submitted (Jan-2005) to the EMBL/GenBank/DBJ databases.
D EMBL: AY914799; AA093441.1; -, mRNA.
D ZFIN: ZDB-GENE-031118-121; zgc:63601.
D InterPro: IPR004944; EGFR_L.
D InterPro: IPR006211; Furin-like.
D InterPro: IPR006212; Furin-repeat.
D InterPro: IPR00719; Prot_Kinase.
D InterPro: IPR002290; Ser_Thr_Kinase.
D InterPro: IPR001245; Tyr_Kinase.
D InterPro: IPR008266; Tyr_Kinase_AS.
D Pfam: PF00057; Furin-like; 1.
D Pfam: PF00069; Kinase; 1.
D Pfam: PF01030; Recep_L_domain; 2.
D PRINTS: PR00109; TYRKINASE.
D ProDom: PD000001; Prot_Kinase; 1.
D SMART: SMO0261; FU; 3.
D SMART: SMO0220; S_TKC; 1.
D SMART: SMO0219; Tyrc; 1.
D PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
D PROSITE: PS00101; PROTEIN_KINASE_DOM; 1.
D PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
D DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 1275 AA; 14082 MW; 67462BAF1B68492F CRC64;

Query Match 51.6%; Score 3514.5; DB 2; Length 1275;
Beet Local Similarity 54.9%; Pred. No. 1.8e-170; Indels 129; Gaps 28;
Matches 721; Conservative 148; Mismatches 315;

QY 9 WGLTALPLPGAASQVCTGTDMKRLPASPETHLMDRLHYOGGVQGNILELYPTN 68
D 11 WVLTLGLTATAGREVLCTGTDMKRLPSLENYEMRLTYTGVQVHGNLEITHLQGN 70
QY 69 ASLSFLQDIQVQGVVLIANQVQVPLQRLRYVGTQLFENYALVALDNDPLANTTP 128
D 71 PDLSEFLQDIQVQGVVLIANQVQVPLQRLRYVGTQLFENYALVALDNDPLANTTP 123
QY 129 VTGASPGGLRELOLRLTEILKGGVLIQORNPOLCYQDTITLMDIPIKNNQALTLIDTNR 188
D 124 -SSQAGLRELRRLSLTEILKGGVLIQORNPOLCYQDTITLMDIPIKNNQALTLIDTNR 182
QY 189 SRACHPCSPMKSGKSCWGESSEDCSLTRTYCAGGACACKPLPTDCHGCAACTGPK 248
D 183 N--CPRGSSACKSGGSCWGEKQDQCTLTSSVMSGSCSKCKPKPSDCHVQCAACTGPK 240
QY 249 HSDCLACHFNHSGICELHCPALVTYNTDTESNPNPEGRKTYFGASCYTAQPYNTLSDY 308
D 241 DSDCLACHFNHSGICELHCPALVTYNTDTESNPNPEGRKTYFGASCYTAQPYNTLSDY 300
QY 309 GSCTLVCPILHNOEYLA--EDG--TORCEKSKPCARVCYVGMEHLREYRAVTSANIOF 364
D 301 -ACTNVCPRKANKVEYIVSEPDDQETQCKEGBCEKVCYVGLGMGLQGVSVNSNIGMF 359
QY 365 AGCKKIFGSLAFLPESFDGDPASNTAPLOPELOQVFTLEITGLYLTISAMPDLSPLDSV 424
D 360 TGCEKIVGSLAFLPESFDGDPASNTAPLOPELOQVFTLEITGLYLTISAMPDLSPLDSV 419
QY 425 FQNLQVIRGTLHNGANSLTLQGLGISWLGRLSRLBELSGALITHNTHLCFVHTVPMQD 484
D 420 FENLQVIRGTLHNGANSLTLQGLGISWLGRLSRLBELSGALITHNTHLCFVHTVPMQD 479
QY 485 LFRNPHQ--LLHTNANRPEDSCVGGGLACHOLCARGHGCMGPGPTQCVNCSQPLRGQECVE 542
D 480 LL-HFTQPNLISNNKQOQCTVSEBGLCDPLGDSGCMGPGPTQCVNCSQPLRGQECVE 538
QY 543 BCRVLQGLPREFYVNAHRLCPHPECOPNGSVTCFGEADQCVACAHYKDPFCVACPS 602
D 539 LCNVLHGSVRPEBGFNCVPCHPBECRPINGTASCTGPPDQCTDCMHRQDGGVCTVERCP 598
QY 603 GVKPDLSTMPYWKPPDEEGACQPCPINCTHSVDLDDKCPAEGORASPLTSTISAVVGL 662
D 599 GVKER--OHTWVKYSNATGHCLPCTINCTVSC-PLDDRGCPYQKTXGIGTVAITVGVGL 655

```



```

Qy 663 LVVVGAVFGLIKRRQOKIRKRYTMRLLQSTELVEPLTPSGAMPNQMILKETELRK 722
Dy 656 LFIILLALVFLYLRQKHQKQKKETIRRLQHELEVEPLTPSGAMPNQMILKETELRK 715
Qy 723 VKVUGSGFVYVGIWTPDEENYKIPVAIKVLRBNTPSKANKELIDEAAYMAGVSGYV 782
Dy 716 LRVVSGAFVGFVFGIMAPDSENVRIPAIVLRBNTPSKANKELIDEAAYMAGVSGYV 775
Qy 783 SRLIGICLTSTVQVLTQAMPYGCCLLDHYENRGRLSGODLLMCMQIAKMSYLEDVRLV 842
Dy 776 CRLIGICLTSTVQVLTQAMPYGCCLLDHYENRGRLSGODLLMCMQIAKMSYLEDVRLV 835
Qy 843 HRDLAARNVLYKSPNHVKITDPGLARLIDIDETEHADGKVPYIKMALISILRRRFTHQ 902
Dy 836 HRDLAARNVLYKSPNHVKITDPGLARLIDIDETEHADGKVPYIKMALISILRRRFTHQ 895
Qy 903 SDVWSYGVTVWELMTFGAKPYDGIPIAREIPDLLEGEPLPQPCITDVMYIMKCMQID 962
Dy 896 SDVWSYGVTVWELMTFGAKPYDGIPIAREIPDLLEGEPLPQPCITDVMYIMKCMQID 955
Qy 963 SECPREFELYSERPMARDPQRFVVIQNEP-LGPASPLDSTFYRSLT--EDDMGDLVD 1019
Dy 956 PDNRPRFDLVDEFTTMARDPSRYVIVIONEDQMSLSTVDSFPRILMAEGGVKKEFLD 1015
Qy 1020 AEEYLVPOQG--FFCPDPAPAGAGVNHRRSSSTR-----SGGGDI----- 1059
Dy 1016 AEEYLVPOQGIFFMTHGEMRANPSRHHSHRSTQMVVEVDGLPNGRELYSVSMIEGSOY 1075
Qy 1060 -TL-----GLEPSEERAP--RSLAPSEAGSD-VF-DDDLGMGAAGKGLPLTHDPS 1107
Dy 1076 PTLVVGATNANMGWGTQYPLARLSISHSAGQSDSVFLDYVE-----DSCP---PS 1125
Qy 1108 PLQRYSEDPYVPL-----PSETFG-----VYAPLTCSFPQRYVYVQPDVRPQPPREGPL 1157
Dy 1126 SPCKYSKDPIMPNGIDGLDLETDMNVRLSHLPLRGTHQPEYVQ----- 1170
Qy 1158 PAAPDAGATLERPKYLSFGK-----NGVKOVFAFGAVENDEYITPQGGAAPQDHP 1209
Dy 1171 -DWAASERPSLT--PRKASERRFILNGL-----STGNSVENPEYLVPIGSIITPT--- 1215
Qy 1210 PPAASPAADNLVYMDQDPEPBGAPPSPTKG-----TPFAENVEYIGL 1251
Dy 1216 ---SPAPDNPYHDIKAAQAVARVALINGTNRHQPNGFMTPFAENVEYIGL 1264

RESULT 9
Q80Y89 MOUSE PRELIMINARY; PRT; 711 AA.
ID Q80Y89_
AC Q80Y89_
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE V-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
neuro/glioblastoma derived oncogene homolog, isoform 2.
GN Name=Erbb2;
OS Mus musculus (Mouse).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
Hogkin R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
Stephenson M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
Brownstein M.J., Ueda T.B., Yoshizuki S., Carrinci P., Prange C.,
Rana S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Boeck S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski N.I., Skalek U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Maira M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046811; AAH46811.1; -; mRNA.
DR EMBL; BC053078; AAH53078.1; -; mRNA.
DR HSSP; P06994; IN8Y.
DR SMK; Q80Y89; 23-630.
DR Ensembl; ENSMUSG00000062312; Mus musculus.
DR MGI; MGI:95410; Erbb2.
DR GO; GO:0016324; C:apical plasma membrane; IDA.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007507; F:heart development; IMP.
DR GO; GO:0042552; P:myelination; IMP.
DR GO; GO:0007422; P:peripheral nervous system development; IMP.
DR InterPro; IPR00494; EGFR_L.
DR InterPro; IPR06212; Furin repeat.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recept_L_domain; 2.
DR SMART; SM00261; FU; 4.
SQ SEQUENCE 711 AA; 78707 MW; 682B18EB0E71318 CRC64;

Query Match 47.8%; Score 3256.5; DB 2; Length 711;
Best Local Similarity 84.5%; Pred. No. 1.2e-157;
Matches 589; Conservative 40; Mismatches 67; Indels 1; Gaps 1;
Qy 1 METALATCRWGLLALALPPGAASITQVCTGDMKRLPASPEITHLDMRLHYGCCVVGQNTL 60
Dy 1 MELAAKCRMGFLALLSPGAAGTQVCTGDMKRLPASPEITHLDMRLHYGCCVVGQNTL 60
Qy 61 ELTYLPTNASTLSPLQDIQEVQGVYLAHNOVROVPLQRLIRVGTQLPEDNYALALVDNG 120
Dy 61 ELTYLPANASTLSPLQDIQEVQGVYLAHNRVKVPLQRLIRVGTQLPEDNYALALVDNR 120
Qy 121 DPLNN-ITPYTASPGSLARELQSLTEILKGGVLIQGNPOLCYQDWMKDVLRKNNQL 179
Dy 121 DPLNNVTTAARGRPESLREILQSLTEILKGGVLIQGNPOLCYQDWMKDVLRKNNQL 180
Qy 180 ALTLIDNRSRACHPGSCPMKSGSRGWSESEDCSLTRTVACAGSARCKGFLPTDCHEQ 239
Dy 180 APVMDNRSRACHPGSCPKDKNHCKGESPEDCOILGTCITSCACRKGFLPTDCHEQ 240
Qy 240 CAAGCTGPKASDCLAHFNHSGICELHCPALVYNTDTFESMNPBGRYTFGASCVTAC 299
Dy 241 CAAGCTGPKASDCLAHFNHSGICELHCPALVYNTDTFESMNPBGRYTFGASCVTTC 300
Qy 300 PNYTLSTDVSGCTLVGLAHQEVTAENGTQRCCKSPCARVCYGLGMEHLRBRATYSA 359
Dy 301 PNYTLSTEVSSCTLVCPNNQEVTAEDGTQRCCKSPCARVCYGLGMEHLRBRATYSD 360
Qy 360 NIOFACCKKIFGSLATLPSFDDPASNTAPLOPQVETLEITGYLYISAMPDSL 419
Dy 361 NIOFACCKKIFGSLATLPSFDDPASNTAPLOPQVETLEITGYLYISAMPESF 420

```

QY 420 PDLSPFQNLQVIRGRILHNGAVSLTLOGLISWLGRLSRLREISGLALIHNTHLCPVHT 479  
 DB 421 QDLSPFQNLQVIRGRILHNGAVSLTLOGLISWLGRLSRLREISGLALIHNTHLCPVHT 480  
 QY 480 VPMDDLFPNPHQALHTANRPEDECGEGLACHOLCARGHCWPGPTCCVNCOSQFLRGOE 539  
 DB 481 VPMDDLFPNPHQALHTANRPEDECGEGLACHOLCARGHCWPGPTCCVNCOSQFLRGOE 540  
 QY 540 CVEBCRVQLGPREYVNAHCLPCHEPCOPQNGSVTCFEPADQCVACAHYDPEFCVAR 599  
 DB 541 CVEBCRVWKGPREYVNRGHCPCHEPCOPQNGSVTCFEPADQCVACAHYDSSCVAR 600  
 QY 600 CPSSGVKPDLSYMPIMKFPDEBEGACOPCPINCTHSCVDLDDKCCPAEORASPLTISIISAVY 659  
 DB 601 CPSSGVKPDLSYMPIMKFPDEBEGACOPCPINCTHSCVDLDDKCCPAEORASPLTISIISAVY 660  
 QY 660 GILLVVLGVFGLILKRRQOKIRKYMRLLOETEL 696  
 DB 661 GVLFLIIVVIGILIKRRRQOKIRKYMRLLOETEV 697

RESULT 10  
 Q9OX70\_RAT PRELIMINARY; PRT; 1209 AA.  
 ID Q9OX70;  
 AC Q9OX70;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DE Epidermal growth factor receptor.  
 GN Name=Egfr;  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 CC Muridae; Murinae; Rattus.  
 CC NCBI\_TaxID=10116;  
 RN [1]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Fisher; TISSUE=Liver;  
 RX MEDLINE=90258888; PubMed=2342466;  
 RA Petch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,  
 RA Earg H.S.;  
 RT "A truncated, secreted form of the epidermal growth factor receptor is  
 RT encoded by an alternatively spliced transcript in normal rat tissue.";  
 RL Mol. Cell. Biol. 10:2973-2982(1990).  
 RN [2]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Fisher; TISSUE=Liver;  
 RA Guttridge K., Dawson T.L., Earg H.S.;  
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Fisher; TISSUE=Liver;  
 RA Petch L.A.;  
 RL Submitted (NOV-1991) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: M37394; AAF14008.1; -, mRNA.  
 DR PIR: A36325; A36325.  
 DR HSSP: Q9H2C9; 1M17.  
 DR SMR: Q9OX70; 25-525, 27-636.  
 DR Ensembl: ENSRNOG0000004332; Rattus norvegicus.  
 DR RGD: 2543; Egfr.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0005488; F:electron transporter activity; IEA.  
 DR GO: GO:0005006; F:epidermal growth factor receptor activity; IEA.  
 DR GO: GO:0005006; F:iron ion binding; IEA.  
 DR GO: GO:0004872; F:receptor activity; IEA.  
 DR GO: GO:0006118; P:electron transport; IEA.  
 DR GO: GO:0006466; P:protein amino acid phosphorylation; IEA.  
 DR GO: GO:0007165; P:transmembrane receptor protein tyrosine kin. .; IEA.  
 DR InterPro: IPR001450; 4Fe4S\_ferredoxin.  
 DR InterPro: IPR006211; EGRF\_L.  
 DR InterPro: IPR006212; Furin\_repeat.

DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase\_AS.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L domain; 2.  
 DR PRINTS: PR00353; 4FE4S\_FDOXIN.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00261; FU; 4.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW Receptor.  
 SQ SEQUENCE 1209 AA; 134890 MW; 96FE7E6CC1B7773 CRC64;

Query Match 46.5%; Score 3172; DB 2; Length 1209;  
 Best Local Similarity 50.2%; Pred. No. 4,6e-153; Indels 124; Gaps 27;  
 Matches 644; Conservative 166; Mismatches 348;

QY 3 LAALCRMGILLALALPPGA-ASTOVCTGTDMLRLPASPEHLDMLRHLYOGCQVQGNLE 61  
 DB 15 LAALCAAG-----GALBEKVCQGSNLTQGTGTFEDHFLSQRMFNNECVLGNLE 66  
 QY 62 LTYLPTNASTLFLDIDQEVGVLIHANQVQVPLQRLIVRGQLFEDNTALAVLNDG 121  
 DB 67 ITTVGRNVDLSFLKTIQEVAGVLIANTVERIPLENQIIRGNALYENTYALAVLSN-- 124  
 QY 122 PLANTTPYTGASPGGLRELQLRLTEILKGVLIQRNQLCYQDTILWKDIFHKNNQAL 181  
 DB 125 -----YGNKTKGLRELPMKNLOEILIGAVRFSSNPILCMETIQRWDIV-QDVFLSN 175  
 QY 182 TLIDTNRS-RACHPSPWCCKSGRCWESSHDCQSLTRTVCAAGCA-RCKGPLPTCCHEQ 239  
 DB 176 MSMDVQRHLTGCPKCDPSCPNCSGCRGEBNCOKLTKIICAOQCRGRGRSPSCCINQ 235  
 QY 240 CAAGCTGPKSHSDCLAHFNHSGICELHCPALVTYNTDTPESMPNBSGRYTFGASCVTAC 299  
 DB 226 CAAGCTGPRSDCLVCHFRBATCKDPCPLMLVNPPTYQMDVNPBGVTSFGATCVKKG 295  
 QY 300 PNYNISTVGSCTLVPLHNOEVTADGTQRCCKSKCARVCYGLGMEHLREVAVTSA 359  
 DB 296 PNYVVTDHGSCVRACGDYEV--EEDVYKCKKCDGCRKQNGIIGEFKDTLSINAT 354  
 QY 360 NIOEPAGCKKIFGSLAPLPSFPGDPAANTAPLOEOLQVETTELTYGLYISAMPDL 419  
 DB 355 NIKHKYCTALSGDHLIPVAFKGSFRTPLDRLRLIKTYVELITGFLDIQAMPENW 414  
 QY 420 PDLSPFQNLQVIRGRILHNGAVSLTLOGLISWLGRLSRLREISGLALIHNTHLCPVHT 479  
 DB 415 TDLHAFENLEIRGTRKHGQFSLAVGLNITSLGRLKELSDQDVIISGRNLCVANT 474  
 QY 480 VPMDDLFPNPHQALHTANRPEDECGEGLACHOLCARGHCWPGPTCCVNCOSQFLRGOE 539  
 DB 475 INMKLFGTPNQKTKIMNRRABKDCATNHCPNPLCSSEGCWGPPTDCCVSCQNSRRRE 534  
 QY 540 CVEBCRVQLGPREYVNAHCLPCHEPCOPQNGSVTCFEPADQCVACAHYDPEFCVAR 599  
 DB 535 CVDKCNILEGEPREREVENSEICQCHEPLPTMTNLTGCRGPDNICKAHYVGDGHCVKT 594  
 QY 600 CPSSGVKPDLSYMPIMKFPDEBEGACOPCPINCTHSCVDLDDKCCPAEORASPLTISIISAVY 658  
 DB 595 CPSSGMENNTL-VKFPDANNVCHLCHANCTGAGAGLKC--QDEBGRIPISIANGI 651  
 QY 659 VGILLVVLGVFGLILKRRQOKIRKYMRLLOETELVPLTPSGAMPNQQRILKE 717  
 DB 652 VGGLFIVV-VALGIGLFMRBRQLVRKTRRLQERELVPLTPSGEAPQADHRIAKE 710  
 QY 718 TELRNVKVLGSGAFETVYKGIWIPGNGVKIPVAIKTLRENTSPKANEIIDEAVVMAGV 777  
 DB 711 TEFKKIKVLGSGAFETVYKGIWIPGNGVKIPVAIKELRENTSPKANEIIDEAVVMASV 770  
 QY 778 GSPVYSRLIGICLTSTVQLVTQLMPYGLLDHVRNRRGLASQDLINMCQIATKMSYLE 837

Db	771	DNPVCRLLGICLSTVQVLITQMLPFGCLLDVBRHKONITQSYLLNCCVQJAKMNYLE	830
Qy	838	DVRLVHRDLAARNVAVKSPNHVKITDFGLARLLIDETEHYADGGKVIKMMALESILRR	897
Db	831	DRRLVHRDLAARNVAVTKPOHVAKITDFGLAKLLGAEKEHYAHGSKVIKMMALESILRR	890
Qy	898	RPTHOSDVWSYGVTVWELMTFGAKPYDGI PAIREIDLLIEKGRLEPOPICTIDVYMIWVX	957
Db	891	IYTHOSDVWSYGVTVWELMTFGSKPYDGI PAISEISILEKGRLEPOPICTIDVYMIWVX	950
Qy	958	CMWIDSEGRPRRELVSFESRWARPORFVYVQ-NEDIGPASPILSTFYRSILDEDDMDG	1016
Db	951	CMWIDADSRPKRELILFESRKAARPORLFVQGERHMLDSPTDSNFYRAMEEDMD	1010
Qy	1017	LVDAAEYLVPOQGFPCDPBPAPGAGGMVHHRRSSSTRSGGDLTLGLPSEBEARSPDLA	1076
Db	1011	VVDAAEYLVIPQGF-----NSPST-----SRTPLL	1036
Qy	1077	PSEGAGSVDFODGLCMGAKLQSLPTHDPSPLORESDPVVPLESEF--DGYAARLTCS	1134
Db	1037	SSLASNSN-----SSTVACINRNGSRUYVEDAFLQRTSSDPTSVLEDDIDDTFL-----	1086
Qy	1135	PQREXVNOGDVRRPQCPSPREGLPLAARPAAGATL-----ERPXTLSPGKGVGVQDVAFAGG	1189
Db	1087	PVPEYITNQ-----SVP-KRPGASVQNFVYTHNQPLHAPAGDLHAYQ--PHSN	1133
Qy	1190	AVENEYVL-TEQGAAPQPHPEPAPSPAFNLYWDO-----DP-----PERGA	1232
Db	1131	AVSNREYLTNQ-----PTGASGDSALMIQKSGHMSLDNDYQODFFPKKAX	1181
Qy	1233	PPSTTKGTPTANPEYLGLDVP	1254
Db	1182	PNGIFPKG-PTAENAEYLVAVRP	1202

RL Oncogene Res. 1:375-396(1987).  
 RN [12]  
 RP NUCLEOTIDE SEQUENCE OF 1-29.  
 RX MEDLINE=91107677; PubMed=1988448;  
 RA Haley J.D., Waterfield M.D.;  
 RT "Contributory effects of de novo transcription and premature  
 RT transcript termination in the regulation of human epidermal growth  
 RT factor receptor proto-oncogene RNA synthesis."; J.  
 RL J. Biol. Chem. 266:1746-1753(1991).  
 RN [13]  
 RP NUCLEOTIDE SEQUENCE OF 1-29.  
 RX MEDLINE=85270438; PubMed=2991899;  
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;  
 RT "Characterization and sequence of the promoter region of the human  
 RT epidermal growth factor receptor gene."; Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).  
 RN [14]  
 RP NUCLEOTIDE SEQUENCE OF 25-49.  
 RX MEDLINE=84172183; PubMed=6324343;  
 RA Weber W., Gill G.N., Spiesse J.;  
 RT "Production of an epidermal growth factor receptor-related protein."; Science 224:294-297(1984).  
 RL [15]  
 RP PROTEIN SEQUENCE OF 540.  
 RA Kohda D.;  
 RL Submitted (SEP-1997) to Swiss-Prot.  
 RN [16]  
 RP PROTEIN SEQUENCE OF 687-705; 986-998; 1000-1023; 1026-1030 AND  
 RP 1068-1077, AND PHOSPHORYLATION SITES THR-693; SER-695; SER-1070 AND  
 RP SER-1071.  
 RX MEDLINE=88330814; PubMed=3138233;  
 RA Heisermann G.J., Gill G.N.;  
 RT "Epidermal growth factor receptor threonine and serine residues  
 RT phosphorylated in vivo."; J. Biol. Chem. 263:13152-13158(1988).  
 RL [17]  
 RP PROTEIN SEQUENCE OF 25-39.  
 RX PubMed=15340161; DOI=10.1110/ps.04682504;  
 RA Zhang Z., Hentzel W.J.;  
 RT "Signal peptide prediction based on analysis of experimentally  
 RT verified cleavage sites."; Protein Sci. 13:2819-2824(2004).  
 RL [18]  
 RP PROTEIN SEQUENCE OF 740-744 AND 746-747.  
 RX MEDLINE=85182650; PubMed=2985580;  
 RA Russo M.W., Lukae T.J., Cohen S., Staros J.V.;  
 RT "Identification of residues in the nucleotide binding site of the  
 RT epidermal growth factor receptor/kinase."; J. Biol. Chem. 260:5205-5208(1985).  
 RN [19]  
 RP RECEPTOR ACTIVITY.  
 RX MEDLINE=84191554; PubMed=6325948;  
 RA Micczkowski B., Mosig G., Cohen S.;  
 RT "ATP-stimulated interaction between epidermal growth factor receptor  
 RT and supercoiled DNA."; Nature 309:270-273(1984).  
 RL [20]  
 RP LIGAND BINDING.  
 RX MEDLINE=90003233; PubMed=2790960; DOI=10.1016/0092-8674(89)90867-2;  
 RA Chen W.S., Lazar C.S., Lund K.A., Welsh J.B., Chang C.P., Walton G.M.,  
 RA Der C.J., Wiley H.S., Gill G.N., Rosenfeld M.G.;  
 RT "Functional independence of the epidermal growth factor receptor from  
 RT a domain required for ligand-induced internalization and calcium  
 RT regulation."; Cell 59:33-43(1989).  
 RL [21]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=89278137; PubMed=2543678;  
 RA Margolis B.L., Lax I., Kriz R., Dombalagian M., Honnegger A.M.,  
 RA Hawk R., Givol D., Ulrich A., Schlesinger J.;  
 RT "All autophosphorylation sites of epidermal growth factor (EGF)  
 RT receptor and HER2/neu are located in their carboxyl-terminal tails.  
 RT Identification of a novel site in EGF receptor.";

RL J. Biol. Chem. 264:10667-10671(1989).  
 RN [22]  
 RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND  
 RP ASN-528.  
 RX MEDLINE=96398132; PubMed=8962717;  
 RA Smith K.D., Davies M.J., Bailey D., Repoué D.V., Hounsell E.F.;  
 RT "Analysis of the glycosylation patterns of the extracellular domain of  
 RT the epidermal growth factor receptor expressed in Chinese hamster  
 RT ovary fibroblasts."; Growth Factors 13:121-132(1996).  
 RN [23]  
 RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND  
 RP ASN-603.  
 RX MEDLINE=20198209; PubMed=10731668;  
 RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;  
 RT "Characterization of the N-oligosaccharides attached to the atypical  
 RT Asn-X-Cys sequence of recombinant human epidermal growth factor  
 RT receptor."; J. Biochem. 127:65-72(2000).  
 RL  
 Query Match 46.5%; Score 3167; DB 1; Length 1210;  
 Best Local Similarity 49.8%; Pred. No. 8.3e-153; Matches 632; Conservative 176; Mismatches 346; Indels 116; Gaps 22;  
 11 LIALALPPGAA--STOVCTGTMKRLRPA SPETHLMDLRHLYOGCQVQGNILETYLPTN 68  
 14 LIALCPASRALAEKRVQGTSMKLTGCTEDHFLSLQRMFNCEVVLGNILETYVQRN 73  
 69 ASLSFLQDIOEVQGYVLIHANOVRVPLQRIYRGVLTGTFEDNVALVLDNGDPLANTTP 128  
 74 YDLSFLKTIQEVAGVLIATLNTVERIPLENLQIIRGNMYENSVYALVATVND----- 126  
 129 VTGASPGSLRELQRLSLTEILKGGVLQRIHQOLCYODITLWKDIPHKXNQALTLIDTNR 188  
 127 ---ANKGLKELPMNLQELIHGAVRSNNPALCVESIQNRDIYSSPLSMSSMDPQNH 183  
 189 SRACHPCSPMCKSGRCWGESSEDDCSLTRVCAGGCA-RCKGRLPTDCHEGCAAGCTGP 247  
 184 LGSCKKDPSPCNCSWGAGENQKLTIIQAQCSGRCKGKSPSCCHNCAAGCTGP 243  
 248 KHSDDLACLFHNSGICELHCPALVTYVTFPESHNPNEGRTYTFGASCYVACPYVYSTD 307  
 244 RESDCLVRKFRDEATCKDTCPLMLVNPITYQMDVNPBGXYSFPAATCVKCKCPRYVYTD 303  
 308 VGSCTLVPLNNOETLAEFGORCEKSKPCARVCYVYGMEHLRVRVATSNANIEFAGC 367  
 304 HGSCVRACGADSYEM-EEEDGVKCKCKGCPKVCNGIGIBFPOSLSINATNIGHFYKC 362  
 368 KIFGSLALPESFGDPASNTAPLPQEPQLQVFTLEBITGYLYISAMPDLSPLDSVFN 427  
 363 TISGDLHILVAPRGDSFTHTPRLDPOEIDLITKVKETITGFLILQAMPERRTDLHAEN 422  
 428 LQVIRGRIHNGAVSLTQGLGISWLGRLRELGSLALIHNTLCPVHTVPMDOFLR 487  
 423 LEIRGRKQGFSLAVVSLNTISLGRJKEISDGVIIISGNMLCYANTIMKKIKFG 482  
 488 NPHQALLHTANRPDECCVGBGLAQQLCARHGCGPGTQCVNCSQFLRSGQCVBCEVYL 547  
 483 TSGQTKTIIISNGENSCATQVCHALCSPEGCWPEBRDCCSNVSRGHECVDKNLL 542  
 548 QGLPREYVNAHCLPCHPECOPONGSVTCFPGPADQCAVAKYKDPFCVACSPSGVXP 607  
 543 EGEPRFVNEBECIOCHECLPQAMNITCTRGPNICIGAHYIDGPHCVKTCFPAVGMB 602  
 608 LSYMPKFPDEBEGACQPCPINCTHSCTVDLDDKGPABQBPASPLTISIAVVG---ILIV 664  
 603 NNTLL-VMKYADAGVCHLCHPNCYGCYGPGLBGCPTGPKIP--SIATGMVGAALLLLIV 659  
 665 VVLGVVPGIILIKRROQKIRKXTMRLLQETELVBEVLTPSGAMPNOAKRIKETEIRYK 724  
 660 VALIGG--LPMRRRHIVKRLTLRLLOEREIVBELTTPSGAPNALRLIKETEFKIK 716  
 725 VLGSAGFTVYKGIWIPDGENVKIPVAIKVLRENTSPPRANKELIDEAAYVMAGVSPVYSR 784

Db	717	VLGSAGFVYKGLIPSEGEKKIVAIKELEATSPANKKILDEAVYMSVDNPHVCR	776
Oy	785	LLGLISTVQLVLTQIMLPYGYCLLDHVRNRGRGLSGODLLMCMQIAKMSYLEDVRLVHR	844
Db	777	LLGLICTSTVQLITQLMPPGCLLDVYRHKKONISQGYLLNMCVQIAKGNVYIEDRLVHR	836
Oy	845	DIAARNVLVKSNNHYKITDPGLARLLIDIDETYYHADGGKVPYIKMALISILRRRTHQSD	904
Db	837	DLAARNVLVKEPQHVKITDPGLAKLLIGAEKEKEYHAEQGVPIKMALISILHRTYHQSD	896
Oy	905	VMSYGVTVWELMTFGCAKPYDGI.PAREIPDLLEKGRLLPQPICTIDVYIMVCKMIDSE	964
Db	897	VMSYGVTVWELMTFESKPYDGI.PASEISSLTEKGRLLPQPICTIDVYIMVCKMIDAD	956
Oy	965	CRPRRELVSEFSRNARDPQRFVVIQ-NEDLGPASPLDSTFYRSLLDDEMDGLVDAEY	1023
Db	957	SRKPRELLIERSKNARDPQRLVYIQGDBRMHLPEPTBSNFRALMDEBEDMDVDADAEY	1016
Oy	1024	LVPOGGFFCPCDPAPGAGGVVHHRHSSSTRSG3DGLTGLEPSEEDAPPSPLAPSEGAGS	1083
Db	1017	LIPQGGFF-----SSPSTRPTPLSSLSATS	1044
Oy	1064	DVFDGDDLGMGAAGKGLQSLPTHDPSPLOXYSEDPYPLPSET--DGVVAPLTCSPPEEYVN	1144
Db	1043	N-NSTVACIDRNGIASCPIKDSFLQRYSDPTGALTEDSIDDTFL-----PVXEYIN	1094
Oy	1142	QPDVAPQPPSPREGPLPARAPGATL-----ERPRTLSPGKNVGVQVFAFGAVENDEY	1196
Db	1095	Q-----SVF-KRPAGSVQNPYVYINQPLNPAPSRPHQD--PHSTAVNGEY	1138
Oy	1197	L-TPQGAAPQPHPPAPFSPAFDNLVYWDQ-----DP-----PERGAPSTTEKG	1233
Db	1139	LVTVQ-----PTCVNSTPDSFAHQAQKSHQISLDNPDYQDFFPKAEKPGVIFKG	1185
Oy	1240	TPTANPEYVL 1249	
Db	1190	S-TANNAEYVL 1198	
RESULT 12			
Q8MIL6_PIG			
ID	Q8MIL6_PIG	PRELIMINARY;	PRT; 1209 AA.
AC	Q8MIL6;		
DT	01-OCT-2002 (Tremblrel. 22, Created)		
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)		
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)		
DE	Epidermal growth factor receptor.		
OS	Sue scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euteria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;		
OC	Sus.		
NC	NCBI_TaxID=9823;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Kim J.G., Vallet J.L., Nonneman D., Christenson R.K.;		
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AY117054; AAM7472.1; -, mRNA.		
DR	HSSP: Q9HC9; 1M17.		
DR	SMR; Q8MIL6; 25-525, 27-636.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.		
DR	GO; GO:0004487; F:receptor activity; IEA.		
DR	GO; GO:0006466; F:protein amino acid phosphorylation; IEA.		
DR	GO; GO:0007169; F:transmembrane receptor protein tyrosine kin. .; IEA.		
DR	InterPro; IPR00494; EGFR_L.		
DR	InterPro; IPR006211; Furin-like.		
DR	InterPro; IPR006212; Furin repeat.		
DR	InterPro; IPR00719; Prot_Kinase.		
DR	InterPro; IPR001245; Tyr_kinase.		
DR	InterPro; IPR008266; Tyr_kinase_AS.		
DR	Pfam; PF00757; Furin-like; 1.		

	Query Match	Best Local Match	Similarity	Score	3152.5;	DB 2;	Length	1209;
				49.7%;	Pred. No. 4.6e-152;			
					Matches 632;	Conservative 178;	Mismatches 344;	Indels 117; Gaps 23
Dr	Pfam, PF01030, Recep_L domain, 2.							
Dr	PR00109; TYRKINASE.							
Dr	Prodcm; PD000001; Prot_Kinase, 1.							
Dr	SMART; SM00261; FU, 5.							
Dr	SMART; SM00219; TyKc, 1.							
Dr	PROSITE; PS00107; PROTEIN KINASE ATP, 1.							
Dr	PROSITE; PS0011; PROTEIN KINASE DOM, 1.							
Dr	PROSITE; PS00109; PROTEIN_KINASE_TYR, 1.							
Kw	Receptor.							
Seq	SEQUENCE	1209 AA;	133531 NM;	268B3FB1E13690F	CR664;			
Qy	12	LLALL-----PGAASTOVCTGTDMLRLPASBETHIDMLRHLYOGCOVYQGNLEITYL	65					
Db	11	LLALLAHFQASPALBEKKVQCGISNKLTYGTFEDHFLSQRMFNNECVLLGNLEITYM	70					
Qy	66	PTNASSLSFLODIQEYGVYLIANHQVQVPLQRLRYRGTOLEFEDNYLAAVLNDGPLN	125					
Db	71	QNSYNSLSFLKIQEAVGVYLIANTVKEIKPLENIQIIRGNLYENTHALAVLSN-----	124					
Qy	126	TTPTVTGASPGSLREQLRSLEILLKGGVLIQRNQLCYQDTILMKDIIFHKNNQLALLID	185					
Db	125	---YGANKTGLRELPMNLQELLQGAFRFSNNPLCAESIQRMDYNSPLSNMSDF	180					
Qy	186	TNRSRACHPCSPMKSGSRCMWSESEDCOSLTRVYACAGCA-RCKGPLPTDCHEQCAAGC	244					
Db	181	QSGQSGCPKCPDGLGNSCMWAGKENGQKLTKVCAQCSGRCHGRSPDCCHNQCAAGC	240					
Qy	245	TGPKHSDCLAFHNHSGICEHLCPALVYNTDTFESMPNDEGRYTFGASCYTPACPNYL	304					
Db	241	TGPRESDCLVGRFRDEATCKDTCPPLMLYNPTTYQMDVNLGKYSFGATCVKCKCPRYV	300					
Qy	305	STDVSGCLIVLEPLNQEYTLADEGQROCKGKPCARVYCYGLGMELREAVATSNIOEF	364					
Db	301	VTHDGSCTPRASSDSYEV-EEQGYRCKCKCGPGCKVCMNGSIGEFKDTLSINATNIHF	359					
Qy	365	AGCKKIFGSLAFLPESFDGPASNTAPLOPEQLQVFTLEETITGYLYISNAPDSLPLDSV	424					
Db	360	RNCTISIGDHLILVAIRGDSFTRTPLRDELDILKTKVITIGELLIQANPEKRTGLHA	419					
Qy	425	FQNLQVIRGRILHNGAVSLTQGLGISWLGSLRLSELGSLALIHNTHLCEHTVPMQ	484					
Db	420	FENLEIIRGRKTHQGGESLAVVGLDISLGRSLKEISDGVIVISGNNLTCYANTISWK	479					
Qy	485	LEFRPHQALLHTARPEDECEYBGLAQHQCARHGMQPGTQVCNCSQPLRGQCEVBC	544					
Db	480	LEGTASQTKLIINRSEKCECAWHICNPLCSSSGCCWPEPRDCSCNFRGKECEVKC	539					
Qy	545	RYLQGLPREYVNAHCLPCHPECOFONGSVYCFPEADQCYACAHHYKDPFCVARSPEGV	604					
Db	540	NVLGEEREREPEANLECYQCHPECLPQAKNYTCMGRPDSCYRCAHYIDGHCIVKTCPAGI	599					
Qy	605	KPDLSTYPIWKFPEDEGAQCPICNTGHSQYDLDDKCGPABORASPLTSTISAVY-GILL	663					
Db	600	AGENSTL-IMKFAIDANVHCHLCHPNTCTYCGVGLGCAVDRPKIP--SIATGIVGILL	656					
Qy	664	VVVLGVVFGILLIKRQCKIRKTYMRRLLOELVELVEPLTPSGAMNQOAMRLIKETELRV	723					
Db	657	AVVALTGGLFLRRR-HIVRKRTLRRLQERELVEPLTPSGEAPNQALRLIKETEFKCV	715					
Qy	724	KVLGSGAFGTYKGIWIPDGENVYKIPVALIKLRENTSPKAKELIDEAYVWAGVSPYVS	783					
Db	716	KVLGSGAFGTYKGIWIPGEKVKIPVALIKLRERATSPKAKELIDEAYVWASVDNPFVC	775					
Qy	784	RLIGICTSTVQVLTQMLPYGCLLDHYRENRGRIGSODLLMWCMQIAKMSYLEVDLVH	843					
Db	776	RLIGICTSTVQVLTQMLPFGCLLDYREHNDNGSQGLLWMCQIAKMGNYLREDBRLVH	835					
Qy	844	RDLAARNVLYKSPNVVKITDFGLARLLIDETEHYADGKVPITWMALESTLRRRAFTQS	903					

Db 836 RDLAARNVLTQHVKITDPCGLAKLGAEEKEVHAEGKGVPIKMLALESLHRYVTHQS 895  
 Qy 904 DWASVGVWVWELMTGAKRYDGIIPAREIDLEKGRLEPQPCITDVMVWVCKMIDS 963  
 Db 896 DWASVGVWVWELMTGSKPYDGIIPASEISTVLEKGRLEPQPCITDVMVWVCKMIDA 955  
 Qy 964 ECRPFRELVSFSSRMARDPQRFVVIQ-NEDLGASPLDSTVRSILLEDMDGLVDABE 1022  
 Db 956 DSRPFRELIIETFSKARDPQRFVVIQSGDERHMLPPTDSNRYALMDEDEMDVDYDABE 1015  
 Qy 1023 YLVPOQGFPCDPAPAGAGMWHHRHSSSTRSGGDLTLGLEPSEEARPSPLABEGAG 1082  
 Db 1016 YLVPOQGFPC-HSPATSRPTLSSLSATSSST-----PAVACVDRNG-- 1054  
 Qy 1063 SVVFPQDGLMGAAKGLQSLPHTDPSRLQYSEDPTVPLPSET--DGYVAPLTCSPQPEV 1140  
 Db 1055 -----QSYPLKEDSFQRYSSDPFGALTEDSLDFTL-----PAPEV 1092  
 Qy 1141 NQPDVRFPPSPREGPLPAPAPAGATL-----ERPKITSPGKNGVYKQVFAFGAVENPE 1195  
 Db 1093 NO-----SVF-KRPAGSVQNPVYHNPISAAFGDPRHYQN--SHSNVAGNPE 1136  
 Qy 1196 YL-TEQGGAQPQHPHPAPSPAFDMLYVWQ-----DP-----PERGAPPSTPK 1238  
 Db 1137 YLNTDR-----PACINGGLDGPAPFAWQTGSHQINDNPDYQQAFFPKCAKSNIGICK 1187  
 Qy 1239 GTPFTANPEYL 1249  
 Db 1188 G-PAENAEYLL 1197

## RESULT 13

EGFR\_MOUSE

ID\_EGFR\_MOUSE STANDARD; PRT; 1210 AA.

AC Q01279;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Epidermal growth factor receptor precursor (EC 2.7.1.112).

GN Name=Egfr;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridea; Muridae; Murinae; Mus.

NC NCBT\_Taxid=10090;

RN NUCLEOTIDE SEQUENCE.

RC STRAIN-BALB/c; TISSUE=Liver;

RX MEDLINE=93026370; PubMed=1408137;

RA Aviyil A., Skorecki K., Yayon A., Givol D.;

RT "Promoter region of the murine fibroblast growth factor receptor 2

(bek/KGFR) gene.";

RL Oncogene 7:1957-1962(1992).

RN NUCLEOTIDE SEQUENCE.

RC STRAIN-BALB/c; and CD-1; TISSUE=Decidua, and Liver;

RX MEDLINE=93126380; PubMed=7678348;

RA Paria B.C., Das S.K., Andrews G.K., Day S.K.;

RT "Expression of the epidermal growth factor receptor gene is regulated

in mouse blastocysts during delayed implantation.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).

RN NUCLEOTIDE SEQUENCE.

RC STRAIN-BALB/c; TISSUE=Liver;

RX Hibbs M.L.;

RL Submitted (Apr-1994) to the EMBL/Genbank/DBJ databases.

RN NUCLEOTIDE SEQUENCE.

RC STRAIN-B6/C3; TISSUE=Liver;

RX MEDLINE=94170986; PubMed=815255;

RA Luetjens N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S.,

RA Jenkins N.A., Lee D.C.;

RT "The mouse waved-2 phenotype results from a point mutation in the EGF

RT receptor tyrosine kinase.";

RL Genes Dev. 8:399-413(1994).

RN [5]

RP NUCLEOTIDE SEQUENCE OF 1-714.

RC TISSUE=Brain;

RX MEDLINE=91232866; PubMed=2030916;

RA Aviyil A., Lax I., Ullrich A., Schlessinger J., Givol D., Morse B.;

RT "Comparison of EGF receptor sequences as a guide to study the ligand

binding site.";

RL Oncogene 6:673-676(1991).

RN (6)

RP NUCLEOTIDE SEQUENCE OF 969-1117.

RC STRAIN=C3H;

RA Bisinger D.P., Serrero G.;

RL Submitted (Jun-1992) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: The EGF receptor mediates the biological signal of EGF,

and also of TGF-alpha, amphiregulin, heparin-binding EGF, Gp30 and

vaccinia virus growth factor (By similarity).

CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein

tyrosine phosphate.

CC -1- SUBUNIT: Binds RPK1. Part of a complex with ERBB2 and either

PIK3C2A or PIK3C2B. The autophosphorylated form interacts with

PIK3C2B, maybe indirectly (By similarity).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to

dimerization, internalization of the EGF-receptor complex,

induction of the tyrosine kinase activity, stimulation of cell DNA

synthesis, and cell proliferation.

CC -1- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor

subfamily.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use as long as its content is in no way modified and this statement is not

removed.

CC EMBL, X78987; CAA55857.1; -; mRNA.

CC EMBL, U03425; AA117899.1; -; mRNA.

CC EMBL, X59698; CAA42219.1; -; mRNA.

CC EMBL, L06864; AAAS3029.1; -; mRNA.

CC EMBL, Z12608; CAA78249.1; -; mRNA.

CC PIR, A53183; A53183.

CC HSP, O9H2C9; 1M17.

CC SMR, Q01279; 25-525, 27-636.

CC Ensembl; ENSMUSG0000020122; Mus musculus.

CC MGI, MGI:95294; Egfr.

CC GO, GO:0016323; C:basolateral plasma membrane; IDA.

CC GO, GO:0030139; C:cytosolic vesicle; IDA.

CC GO, GO:0016021; C:integral to membrane; TAS.

CC GO, GO:0005622; C:intracellular; IDA.

CC GO, GO:0016301; F:kinase activity; IDA.

CC GO, GO:0005515; F:protein binding; IPI.

CC GO, GO:004871; F:signal transducer activity; IDA.

CC GO, GO:0050730; P:regulation of peptidyl-tyrosine phosphorylation; IMP.

CC GO, GO:0007165; P:signal transduction; IDA.

CC InterPro, IPR001450; 4Fe4S\_Fe\_S\_bd.

CC InterPro, IPR000494; Egfr\_L.

CC InterPro, IPR006211; Furin-like.

CC InterPro, IPR000719; Prot kinase.

CC InterPro, IPR001245; Tyr\_pkinase.

CC InterPro, IPR008266; Tyr\_pkinase\_AS.

CC Pfam, PF00757; Furin-like; 1.

CC Pfam, PF01030; Recep\_L\_domain; 2.

CC PRINTS, PR00109; TYRKINASE.

CC PRINTS, PR00353; 4Fe4SFRDOXIN.

CC PRODOM, PD000001; Prot kinase; 1.

CC PROSITE, PS00107; PROTEIN\_KINASE\_ATP; 1.

CC PROSITE, PS50011; PROTEIN\_KINASE\_DOM; 1.

CC PROSITE, PS00109; PROTEIN\_KINASE\_TYR; 1.

CC ATP-binding; Glycoprotein; Kinase; Nucleotide-binding;

Phosphorylation; Receptor; Repeat; Signal; Transferase; Transmembrane;

Tyrosine-protein kinase.

KW

FT	SIGNAL	1	24	Potential.
FT	CHAIN	25	1210	Epidermal growth factor receptor.
FT	TOPO_DOM	25	647	Extracellular (Potential).
FT	TRANSMEM	648	670	Potential.
FT	TOPO_DOM	671	1210	Cytoplasmic (Potential).
FT	REPEAT	75	300	Approximate.
FT	REPEAT	390	600	Approximate.
FT	DOMAIN	714	981	Protein kinase.
FT	NP_BIND	720	728	ATP (By similarity).
FT	COMBIDAS	1028	1071	Ser-rich.
FT	ACT_SITE	839	839	By similarity.
FT	BINDING	747	747	ATP (By similarity).
FT	SITE	1018	1018	Important for interaction with PIK3CB (By similarity).
FT	MOD_RES	680	680	Phosphothreonine (by PKC) (By similarity).
FT	MOD_RES	1092	1092	Phosphotyrosine (by autocatalysis) (By similarity).
FT	MOD_RES	1110	1110	Phosphotyrosine (by autocatalysis) (By similarity).
FT	MOD_RES	1172	1172	Phosphotyrosine (by autocatalysis) (By similarity).
FT	MOD_RES	1197	1197	Phosphotyrosine (by autocatalysis) (By similarity).
FT	CARBOHYD	128	128	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	175	175	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	196	196	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	352	352	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	413	413	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	444	444	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	528	528	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	568	568	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	603	603	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	623	623	N-linked (GlcNAc. . .) (Potential).
FT	DISULFID	190	199	By similarity.
FT	DISULFID	194	207	By similarity.
FT	DISULFID	215	223	By similarity.
FT	DISULFID	219	231	By similarity.
FT	DISULFID	232	240	By similarity.
FT	DISULFID	236	248	By similarity.
FT	DISULFID	251	260	By similarity.
FT	DISULFID	264	291	By similarity.
FT	DISULFID	295	307	By similarity.
FT	DISULFID	311	326	By similarity.
FT	DISULFID	329	333	By similarity.
FT	DISULFID	506	515	By similarity.
FT	DISULFID	510	523	By similarity.
FT	DISULFID	526	535	By similarity.
FT	DISULFID	539	555	By similarity.
FT	DISULFID	558	571	By similarity.
FT	DISULFID	562	579	By similarity.
FT	DISULFID	582	591	By similarity.
FT	DISULFID	595	617	By similarity.
FT	DISULFID	620	628	By similarity.
FT	DISULFID	624	636	By similarity.
FT	CONFLICT	19	19	C -> S (in Ref. 2).
FT	CONFLICT	539	539	C -> W (in Ref. 5).
FT	CONFLICT	991	991	L -> F (in Ref. 4).
FT	CONFLICT	1116	1117	HP -> DR (in Ref. 6).
SO	SEQUENCE	1210 AA;	134853 MW;	690E20D46FD22D2F5 CRC64;
Query Match				
Best Local Similarity 49.7%; Score 3146; DB 1; Length 1210;				
Matches 635; Conservative 169; Mismatches 353; Indels 120; Gaps 24;				
QY	11	LLALLPFGAA--STQVCTGDTMKRLPASPETHLDMLRLHYGCGVQVQGNLETLVLTN	68	
DB	14	LTLTCAAGALAEKKVCGSTSNRLTQGTPEDFHLSLQRMYNCEVVLGNLEITTVGRN	73	
QY	69	ASLSFLQDIOEVGVVLTAFHNOVROVPLORLRIYVGTOLFEDNVYALAVLDNGDPLANTTP	128	
DB	74	YDLSPKTIQEVAVGVVLTALNTVERIPLENTQIINGNALYENTYALILSN-----	124	

QY	129	VTGASPGGLARELQRLSRLTEILKGVVLQRPOLCYODTIIMKDI---FHKNOLATLTI	184	
DB	125	-YGNTRGRLRELPRNLOEILIGAVRSPNNPILCNMDITQWRDIVQNVFMNMSMDL---	180	
QY	185	DTNRSRACHPSPKCKSRCKGSESDCOQLTRVVCAGGCA-RKGPLPTCCCHQCGAG	243	
DB	181	-QSHSPSCPKDPSCPNKSGWGGSEENCQKLTILCAQQCSHRCRKGSPSCCHQCGAG	239	
QY	244	CTGPKHSDCLACLFHFNHSGICELHCPALVTYNTPTFESMPNDEGRYTFGASCVTACPYNY	303	
DB	240	CTGPRESDCLVQCFQFQDEATKQCTPLMLNPTTYQMDVNPBKYSFGATCVKCKPRNY	299	
QY	304	LSTVSGCTVCLPHNDEVTAEDETQRCCKSKFCARVCYGLGHEHLREVAVTSANIQE	363	
DB	300	VTDHSGCVRAKCGPDYEV--EEDGIRKCKKDCGCRKVCNIGIGEGFDTISINATNIKH	358	
QY	364	FAGCKKIFGLSLAPLPESPDGDPASNTAPLOEQLQVFTLEITGYLVISAMPSPPLTS	423	
DB	359	FKYCTALSGDLHLIPVAFKDSFTPTPLDRELEILKTEITGFLLIQAMPDWDLH	418	
QY	424	VFQNLQVIRGRILHNGAVSLTLQGLISWGLRLSLRLSLGSLALIHNTHLCPVHTVPM	483	
DB	419	AFENLEIIRGRTKHQGFSLAVGLNITSLGLRLKELISDGVIIISGRNLCVANTIMWK	478	
QY	484	QLFNPHQALHTANRPDESCVSEGLACHQCLANRHCCKGPEPTQCVNCSQFLRQGEVSE	543	
DB	479	KLFTPTMQTKIMNNRRAEKDCAVNHVNCPLCSSEBGCWPBPBCVCSQNSRGECEK	538	
QY	544	CRVLQGLPREVYNNRHLCPHRECOPONGSVTCGPEADOCVACAHYDPFPCVACRPSG	603	
DB	539	CNIEGEBREVENSECTQCHPECLPQAMNITCTGRGPDNCTIQAHYIDGHCYKTCAG	598	
QY	604	VKPLSYMPIKPFDEBACQPCPINCDSVDDLDDKCPAPORASPLTISAVVGLL	663	
DB	599	IMGENNTL-VWKYADANNVCHLCHANCTYGCAGGLQCEVWSPGPKPSTATGVGLL	657	
QY	664	VVVLGVVFGI-LIRROQKIRKTYMRLLQETELVEPLTBSGAMPNOQMLKTELRK	722	
DB	658	FIVV-VALLGIGLFMRRRHIVAKRTLRLQRELEVEPLTBSGEBNQAHLRLKTEPEK	716	
QY	723	VKVLSGAFGVVYGIWIPQSENVKIPVAILVLRNTPSKANKELIDAYYMAVGSPYV	782	
DB	717	IKVLSGAFGVVYGLWIPSEKVKIIPVALKELBEANSPKANKELIDAYYMAVDNHHV	776	
QY	783	SRLGICLTSTVQQLVTOQLMPYGLLDHYENRGRIGSQDLLMCMQIAKGSYLEDVLY	842	
DB	777	CRLLGICLTSTVQQLVTOQLMPYGLLDYREHKDNIIGSYLLMWCQIAKGNVLEDRLV	836	
QY	843	HRDLAARNVLYKSPNHYKITDPGLARLLDDETYHADGKVPYIKMALLESILRRFTHQ	902	
DB	837	HRDLAARNVLYKTPQHYKITDPGLAKLGAEBEYHAEKGVPYIKMALLESILHRIYHQ	896	
QY	903	SDVWSYGVTWELMTFGAKPYDGI PAEREIPLLEKGERLPORPCTTDVYIMVYKMWID	962	
DB	897	SDVWSYGVTWELMTFGSKPYDGI PASDISILKEGERLPORPCTTDVYIMVYKMWID	956	
QY	963	SECPREFRELYSESRMARDPQRFVVIQ-NBDLGPASPLDSTFYRSLDEMDGVDYAE	1021	
DB	957	ADSRPKRELELLEBSKMARDPQRYLVIGDGRMKLPSTDSNFPRLAMDESDMEDVDAD	1016	
QY	1022	EYLVPOQGFPCPDPAFGAGVHHRHNSSTRSGGGLTGLLBSSEBAFSPPLAPSGA	1081	
DB	1017	EYLVPOQGF-----NSPST-----SRTPLSLISLSA	1042	
QY	1082	GSDVFDGLDGMGAAGLQSLPTHPSPLORESDEPTVLPSET--DGVAVPLTSPPOPEY	1139	
DB	1043	TSN-----NSTVACINRNGSCVYKEDATLQRYSSPTGAVTENDINDDAVL-----PVREY	1092	
QY	1140	VNCPDVVPQPPSPREGPLPAARPAAGATL-----ERPKTLSPGKGVVQDVAFGAVENP	1194	
DB	1093	VNQ-----SWP-KRPAGSVQNPYYVHNQPLHAPAGRDLYQN--PHSAVQNP	1136	
QY	1195	EYL-TPGGAAPQHPHPAPSPAFDNLTYWDQ-----DP-----BERGAPSTF	1237	



Db 1137 EYLTAAQ-----PCTLSGFSNPAIMIQGSHQSLMDPDYQDPPFKETKXGIF 1187  
Qy 1238 KGTPTAENPEYIGLDVP 1254  
Db 1188 KG-PTAENAEYLRVAP 1203  
RESULT 14  
Q5SV88\_MOUSE PRELIMINARY; PRT; 1210 AA.  
ID Q5SV88\_MOUSE PRELIMINARY; PRT; 1210 AA.  
AC Q5SV88;  
DT 01-FEB-2005 (Tremblrel. 29, Created)  
DT 01-FEB-2005 (Tremblrel. 29, Last sequence update)  
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)  
DE Epidermal growth factor receptor.  
GN Name=Bgfr; ORFNames=RP23-295E4.1-001;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Muridae; Murinae; Mus.  
CX NCBI\_Taxid=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Graffham D.;  
RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Matthews L.;  
RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AL645532; CA12498.2; -; Genomic DNA.  
DR EMBL; BX255277; CA135984.1; -; Genomic DNA.  
DR EMBL; BX255277; CA12498.2; JOINED; Genomic DNA.  
DR EMBL; AL645532; CA135984.1; JOINED; Genomic DNA.  
DR SMR; Q5SV88; 25-525, 26-636.  
DR Ensemble; ENSMUSG00000020122; Mus musculus.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.  
DR GO; GO:0004674; F:protein ion binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR GO; GO:0007169; P:citranemembrane receptor protein tyrosine kin. .; IEA.  
DR InterPro; IPR001450; 4Fe4S\_Fe\_S\_bd.  
DR InterPro; IPR000494; EGFR\_L.  
DR InterPro; IPR006212; Furin-like.  
DR InterPro; IPR006212; Furin repeat.  
DR InterPro; IPR000719; Prot\_Kinase.  
DR InterPro; IPR002290; Ser\_Thr\_Kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF01030; Recep\_L\_domain; 2.  
DR PRINTS; PR00353; 4FESGRDOKIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_Kinase; 1.  
DR SMART; SM00261; FU; 4.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW Receptor.  
SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DE2D2F5 CRC64;  
Query Match 46.2%; Score 3146; DB 2; Length 1210;  
Best Local Similarity 49.7%; Pred. No. 9, 8e-152;  
Matches 635; Conservative 169; Mismatches 353; Indels 120; Gaps 24;  
Qy 11 LLLALLPGAA--STGVCTGTDMLRLPASPETHLMDRLHYGGQGVVQGNLELTLPNTN 68

Db 14 LITLALCAAGALAEKKVCGSTSNRLTQGFEDHPLSLQRYMNCSEVLLGNLEITYVORN 73  
Qy 69 ASLSFLDIOIQVQGVVLIANHQVQPLQRLRYVGTGLFEDNYLALINDNDPLANTTP 128  
Db 74 YDLSFLKTIQEVAGVLIALTVERIIPLENQIIRGNALYENTVALILSN----- 124  
Qy 129 VTGASPGGLRELOSLRTEILKGGVLIQNPOLCYODITLWMDI----FHKNNQALATLI 184  
Db 125 -YGNRTKLRELPNRLNLOEILIGAVRSPNNPLCMMDITQMRDIYQNVFMSMSMDL--- 180  
Qy 185 DTRNSRACHPCSPMCKSGRCWGSSEDCQSLTRVYACAGCA-RCKGPLPTDCHEQCAAG 243  
Db 181 -QSHSSCPKCDPCSPNCSWGGEHCQKLTIKICAQCCSHRCGRSPSCCHQCAAG 239  
Qy 244 CTGPKHSCCLACLNHNSGICELHCPALVYNTYTFEEMPRPEGRYTGACVYACPTNY 303  
Db 240 CTGPRESDCLVQCKQDQBATCKDTPPLMLYNPTTYQMDVNPBGKYSFGATCVKCKCPNNY 299  
Qy 304 LSTDVSGCTLVCPRLNQEVTADGTORCEKSKPCARVCYGLGMBHLREVAVTSANIQE 363  
Db 300 VTDHSGCVRAQGPYIEV-BEDGIRKCKKDDGPRKYNCGIGIEFPDITLSINATNKH 358  
Qy 364 FAGCKKIFGSLAPLESFSDPDASNTAPLOEOLQVFETLEITGYLIISAMPSPDLS 423  
Db 359 FVCTAISGDLHILPVAFKDGSPTRTPLDRELEILTKVKEITGFLIIQAMPDWTDLH 418  
Qy 424 VPQNLQVTRGRLINHGAVSLTQIGISWGLRSLRELSGLALIHNTHLCFYNTVWD 483  
Db 419 AFENLEIRGRTHKGQPSLAVVGLNTLSLRSLKESLSDVDVLSIGRNLCYANTIMWK 478  
Qy 484 QLFRRPHOALHTANRPDECEVGEGLACHQLCARHCNGPGPTQCVNCSQFLRGOEYEB 543  
Db 479 KLFGRPNKTIIMNRRAKDCKAVNVHVCNPLCSSSGCGPBRDVCSONSRGECYEB 538  
Qy 544 CRVLOGLEBYVNAHCLPCHPECOPONGYATCFGPRAQDCVACAHYKDPFCVARRCSG 603  
Db 539 CNLIEGEPREVEENSECICQHCPCILPQAMNITTCGRGPDNCICQAHYIDGPHCVKTCBAG 598  
Qy 604 VKPDLSTYPIKFPDEBACQPCPNTCHGCVDDDDKCKPABQASPLTSTISAVVGLL 663  
Db 599 IMGENNTL-VKTYDANNVCHCANCTYGCAGPQLQCEVWSPGKLPJSTATGTVGLL 657  
Qy 664 VVVLGVVGEI-LIKRROCKIRKTYMRRLQETELVEPLTPSGAMPQNAOMELKETEARK 722  
Db 658 FIVV-VALIGLFPERRRIVKRTLRLLQRELVETLPBGEARNQAHLLIKETEKK 716  
Qy 723 VKVLGSGAFGVYKGIWIPDSENVKIPVALIVLRENTSPKANKETLDEAYYVAGVSPYV 782  
Db 717 IKVLGSGAFGVYKGIWIPBSENVKIPVALIETLEAATSPKANKETLDEAYYVAGVSPYV 776  
Qy 783 SRLIGICTSTVOULTQMLPYGCLLDHYRBNRGRGLSGODLNLWCKQIAKMSYLEADVLY 842  
Db 777 CRLLGICTSTVOULTQMLPYGCLLDHYRBNRGRGLSGODLNLWCKQIAKMSYLEADVLY 836  
Qy 843 HRDLAARVNLKSPNHYKITDQGLARLLDIDETETHADGGKVPKIMWALLESILRRRFTHQ 902  
Db 837 HRDLAARVNLKSPNHYKITDQGLARLLDIDETETHADGGKVPKIMWALLESILRRRFTHQ 896  
Qy 903 SDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLEKGERLPORPCTTIDVYIMVYKCMMD 962  
Db 897 SDVMSYGVTVWELMTFGSKPYDGIIPASD19SILKGERLPPRLCTTIDVYIMVYKCMMD 956  
Qy 963 SECRPRFRELVSFSPSKARADPQRVYVLO-NEDLGPASGLDSTFRSLLEDMDKDVDAE 1021  
Db 957 ADSBPFRRELILBPSKARADPQRVYVLOGDERRMLPSPPTDNPFRALMDEEDMDVDVAD 1016  
Qy 1022 EYLVPOQGFPCPDAPAPAGKGVNHRHSSSTRSGGGLTGLBPSSEBAPSPPLAPSGA 1081  
Db 1017 EYLVPOQGF-----NSPST-----STPLLSLSLA 1042  
Qy 1082 GSDVFDGDLGKAAKGLQSLPETHDPSPLQRYSEDPVLPBSET--DGVAVPLTCSPOREY 1139

Db 1043 TSN-----NSTVACINRNGSCRVKEDAFLORYSSDPTGAVTEDNDIDAFL-----PVBEX 1092  
Qy 1140 VNPDVVRQPPSPREGLPAARPAAGATL-----ERPKTLSPGKGVNVDFVAFGAVENP 1194  
Db 1093 VNO-----SVF-KRPACSVONPVYTHNOPLHPAFGRDLHYON--PHSAVGNP 1136  
Qy 1195 EYL-TPQGGAAPQHPHPAPAFSPADNLXYWQO-----DP-----BERGAPSTF 1237  
Db 1137 EYLNATQAQ-----PTCLSSGFNSPALMTQKSGHOMSLDNPDYQODPFPEKTEPNCIGF 1187  
Qy 1238 KGTPTAENPEYLGDDVP 1254  
Db 1188 KG-PTAENAEYLRVAP 1203

RESULT 15  
Q9EP98\_MOUSE PRELIMINARY; PRT; 1210 AA.  
ID Q9EP98\_MOUSE PRELIMINARY; PRT; 1210 AA.  
AC Q9EP98;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-FEB-2005 (Tremblrel. 29, Last annotation update)  
DE Epidermal growth factor receptor isoform 1.  
GN Name=Egfr;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C3H/101, and C57BL/6J;  
RX MEDLINE=21100872; PubMed=11161793; DOI=10.1006/geno.2000.6341;  
RA Beller J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,  
RA Schehl Sincclair C., Pearse R.S., Green P.J., Yee D., Lampland A.L.,  
RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,  
RA Maibhe N.J.;  
RT "Comparative genomic sequence analysis and isolation of human and  
RT mouse alternative Egfr transcripts encoding truncated receptor  
RT isoforms";  
RL Genomics 71:1-20(2001).  
DR EMBL; AF275366; AAC28045.1; -; Genomic\_DNA.  
DR EMBL; AF275367; AAC24386.1; -; mRNA.  
DR EMBL; AF275364; AAC28045.1; JOINED; Genomic\_DNA.  
DR EMBL; AF275365; AAC28045.1; JOINED; Genomic\_DNA.  
DR HSBP; O9H2C9; 1M17.  
DR SRR; O9EP98; 25--525, 27-636.  
DR MGI; MGI:95284; Egfr.  
DR GO; GO:0016323; C:basolateral plasma membrane; IDA.  
DR GO; GO:0030139; C:endocytic vesicle; IDA.  
DR GO; GO:0016021; C:integral to membrane; TAS.  
DR GO; GO:0005622; C:intracellular; IDA.  
DR GO; GO:0016301; F:kinase activity; IDA.  
DR GO; GO:0005515; F:protein binding; IPT.  
DR GO; GO:004871; F:signal transducer activity; IDA.  
DR GO; GO:0050730; P:regulation of peptidyl-tyrosine phosphorylation; IMP.  
DR GO; GO:0007165; P:signal transduction; IDA.  
DR InterPro; IPR001450; 4Fe4S\_ferredoxin.  
DR InterPro; IPR000494; EGFR\_L.  
DR InterPro; IPR006211; Furin-like.  
DR InterPro; IPR006212; Furin\_repeat.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF01030; Recp\_L\_domain; 2.  
DR PRINTS; PR00353; 4FE4SFRDXXIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00261; Fu; 4.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW Receptor.  
SQ SEQUENCE 1210 AA; 134840 MW; 62CD021C9D82E18 CRC64;  
Query Match 46.1%; Score 3143; DB 2; Length 1210;  
Best Similarity 49.7%; Pred. No. 1,4e-151;  
Best Local Similarity 168; Mismatches 354; Indels 120; Gaps 24;  
Matches 635; Conservative 168; Mismatches 354; Indels 120; Gaps 24;  
Qy 11 LLLALPPGAA--STQVCTGTDMLRLPASBETHLDMRLHLXGQGVVQGNILETYLPTN 68  
Db 14 LITLCAAGALHEKRCVQCGTSNRLTQGTEDHFLSLQRYNNCEVLLGLNLEITYQRN 73  
Qy 69 ASISFLQDIOGVGYVLIANQVRQVPLQRLRIYRGVQLFEDNALAVLNDGPLNNTTP 128  
Db 74 YDLFLKTIQEVAGVLIANTVERIDPLENLIIRGNALYENTYALALISN----- 124  
Qy 129 VTGASPGGLRELQRLSLTEILKGVLIQRPQLCYQDTILMKD-----FHNQNLALTLI 184  
Db 125 -YGTNRKGLRELPRNQLQELIGAVRSNNPILCNMDITQMRDIVQNVFMSNMNDL--- 180  
Qy 185 DYNRSRACHPCSPKCKSGKSGESSEDCQSLTFTVCAGGCA-RCKGRLPTDCCHQCAAG 243  
Db 181 -QSHPSGCPKDCPSGSCMGSEBNCQKLTIKCAQCSHRKRGSRSDCNCQCAAG 239  
Qy 244 CTGPKHSDCLACLFNHSIGICELHCPALVTYNTDFFESMPNBSGRYTFGASCVTACPYNY 303  
Db 240 CTGRESDCVLCQKFOEATCKDCCPLMLYNPTTYQMDVNEGKYSFGATCVKCPRNY 299  
Qy 304 LSTVGSCTLVCPHANEVYAEADTORCEKSKRCACVCGVGLGMEHLREVAATVSNIOE 363  
Db 300 VYTDHSGCVRAQGDYEV--EEDGIRCKCKDGCRCRVKNGIGIEFDLTLSINATNKKH 358  
Qy 364 FAGCKITFGSLAFPSFSDGPASNTAPLOEQQLQVETLEIRGYLISAMPSPDLPS 423  
Db 359 FKCYTALSGLHLIPVAFKSDSFTRTPLDREILKTYVEIRGFLIQMPDMDLH 418  
Qy 424 VPQNLQVIRGRILHNGVSLTLOGLGISWGLRSRLBSGLALIHNNHLCPVHYVWD 483  
Db 419 AFENLEIRIGRTHQGFSLAVVGLANTISLRLSLKEISDDVLIISGRNLCYANTIMWK 478  
Qy 484 QLFENPQALLHTNRPEDECVSGGLACHQLCAGDHCMGPRPTCVNCSQRLROECVEE 543  
Db 479 KLFTPNQKTKIMNRKAEKCKAVNHCNPICSESGCGMPRDCVSNVSRKRECEK 538  
Qy 544 CRVLAQGPREYVNAHCLPCHPECOPNGSYTCGPBADOCAVAAHYADPFCVACRPSG 603  
Db 539 CNLIEGPRFEVNSSEICQHPBCLPQAMNITCGRBPNDICQCAHYDGHCVKTCGAG 598  
Qy 604 VKPDLSTYPIWKPFDEGACQPCPINCTHSCVDLDDKGCAPBQASPLTSLISAVGILL 663  
Db 599 IMGENNTL-VWKYDANNVCHLCHANCTYGCAGGGLGCEVWPSGPKIPSIATIGVGLL 657  
Qy 664 VVVLGVVFGI-LIKRROOKIRKYMRLLOETELVEPLTPSGAMPNOAMILLKETEIRK 722  
Db 658 FTVV-VALGIGLEFNRRHIVAKKRLRLQREIVEPLTPSGEAPNQHMLILKETEKK 716  
Qy 723 VKVLGSAFGTVYKGIWIPGENYKIPVALIKVLEBNTSPKANKIILDEAYVMAGVSEPVY 782  
Db 717 IKVLGSAFGTVYKGLIIPGEKKIIPVALIKELREASPRANKIILDEAYVMASVDNPHV 776  
Qy 783 SRLIGICLSTVQLVTLQMPYGLLDHVRNRRGLSQDILNMCQIAKMSYLEDVRLV 842  
Db 777 CRLLGICLSTVQLITQLMPEYGLLDVYREHKDNIQSQYLLNMCVQIAKGNVYLEDRLV 836  
Qy 843 HRDLAANVLYKSNHVKITNDFGLARLLDDETYHADGGVPIKMMALSSILRRRTHQ 902  
Db 837 HRDLAANVLYKTPQHKITDFGLAKLGAEBKRYHAEAGGVPIKMMALSSILRRRTHQ 896  
Qy 903 SDVWSYGVYVWELMTFAKPYDGIIPAREIPDLLEKGRLLPOPICTIDVYVIMYKCMNID 962  
Db 897 SDVWSYGVYVWELMTFSSKPYDGIIPADISSILEKGRLLPOPICTIDVYVIMYKCMNID 956  
Qy 963 SECRPRELIVSEFSRMAQDPQRFVVIQ-NEDLGAPASPLDSTFYRSLLEDDMDGLVDAAE 1021

```

Db 957 ADSPKRFRELLEBSKWARDPQRTLVIGSDERMHLPSPTDSNFYRALMDEEDMEDVVDAD 1016
QY 1022 EYLVPOGFFCPCPDAPAGCMVHHHRSSSTRSGGDLTLGLEPSEBEAPRSPAPSEGA 1081
Db 1017 EYLTPOGFF-----NSPST-----SRTPLSSLSA 1042
QY 1082 GSDVFDGDLGMAKGLQSLPTHDSPLQRYSEDPVPLPSET--DGVAFLTCSPOPEY 1139
Db 1043 TSN---NSTVACINRNGSCRVKEDAFLOKYSDDPTGAVTEDNIDDAFL-----FVPEY 1092
QY 1140 VNQPDVRRPQPPPREGRLPAARPAAGATL-----ERPRTLSPGKNGVWDVPAFGGAVENP 1194
Db 1093 VNQ-----SVP-KRPAGSVQNPVYHNPPLHAPGRDLHYQN--PHSNAVGNP 1136
QY 1195 EYL-TPOGGAAPQHPHPAPSPAFDNLXYWDQ-----DP-----PERGAPSTF 1237
Db 1137 EYLNTAQ-----PTCLSSGFNSPALWIOKSHQMSLDNPDYQODFPKXETKPNIGF 1187
QY 1238 KGPTAENPEYLG DVP 1254
Db 1188 KG-PTAENAEYLRVAP 1203

```

Search completed: January 3, 2006, 11:16:43  
 Job time : 267 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2006, 10:55:57 ; Search time 196 Seconds  
(without alignments)  
2813.369 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 METALCRWGLLALLPFGA.....TFKGTPTANPEYLGIDVTV 1255

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 244163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_21:\*  
1: geneeqp19808:\*  
2: geneeqp19908:\*  
3: geneeqp20008:\*  
4: geneeqp20018:\*  
5: geneeqp20028:\*  
6: geneeqp20038:\*  
7: geneeqp20038:\*  
8: geneeqp20048:\*  
9: geneeqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	2 AAW01111	AAW01111 HER-2/neu
2	6815	100.0	1255	2 AAW92406	AAW92406 Human HER
3	6815	100.0	1255	3 AAY84780	AAY84780 Amino aci
4	6815	100.0	1255	3 AAB21198	AAB21198 Human HER
5	6815	100.0	1255	4 AAC68267	AAC68267 HER2/neu
6	6815	100.0	1255	4 AAB85458	AAB85458 Human HER
7	6815	100.0	1255	5 AAE20479	AAE20479 Human HER
8	6815	100.0	1255	5 AAU77114	AAU77114 Human Her
9	6815	100.0	1255	5 AAW51143	AAW51143 Human Her
10	6815	100.0	1255	5 AAB24067	AAB24067 Human Her
11	6815	100.0	1255	6 ABR43687	ABR43687 Human c-e
12	6815	100.0	1255	7 ABR82066	ABR82066 Human Her
13	6815	100.0	1255	7 ADC09593	ADC09593 Her2/Neu
14	6815	100.0	1255	7 ADD25464	ADD25464 Binding d
15	6815	100.0	1255	7 ADB63281	ADB63281 Human Pro
16	6815	100.0	1255	7 ADE76190	ADE76190 Human HER
17	6815	100.0	1255	7 ADP45048	ADP45048 Human Kin
18	6815	100.0	1255	8 ADJ65554	ADJ65554 Her2 prot
19	6815	100.0	1255	8 ADL90083	ADL90083 Human Her
20	6815	100.0	1255	8 ADQ17193	ADQ17193 Human ref
21	6815	100.0	1255	8 ADI55305	ADI55305 OSPF-reia
22	6815	100.0	1255	8 ADPE0108	ADPE0108 Human epi
23	6815	100.0	1255	8 ADU04914	ADU04914 HER-2/neu
24	6815	100.0	1255	9 ADW44355	ADW44355 Human tyr

25	6815	100.0	1255	9 ADW95472	ADW95472 Amino aci
26	6815	100.0	1255	9 AEB87742	AEB87742 Human ERB
27	6809	99.9	1255	9 ADM87390	ADM87390 Human cod
28	6806	99.9	1255	3 AAY92820	AAY92820 Human her
29	6806	99.9	1255	4 AAB60167	AAB60167 HER2 tran
30	6806	99.9	1255	4 AAE12130	AAE12130 Human tyr
31	6806	99.9	1255	5 AAE26349	AAE26349 Human HER
32	6806	99.9	1255	5 AAE26366	AAE26366 Human Her
33	6806	99.9	1255	5 AAU74545	AAU74545 Human HER
34	6806	99.9	1255	6 ABR47447	ABR47447 Breast ca
35	6806	99.9	1255	6 ABR74708	ABR74708 Human Her
36	6806	99.9	1255	6 AAS83390	AAS83390 Human c-e
37	6806	99.9	1255	6 ADA38143	ADA38143 Human erb
38	6806	99.9	1255	7 ADA37255	ADA37255 Human erb
39	6806	99.9	1255	7 ADB67621	ADB67621 Human epi
40	6806	99.9	1255	8 ADH13187	ADH13187 Human mal
41	6806	99.9	1255	8 ADM72831	ADM72831 Human Her
42	6806	99.9	1255	8 ADO20009	ADO20009 Human PRO
43	6806	99.9	1255	8 ADQ29700	ADQ29700 Human col
44	6806	99.9	1255	8 ABW81864	ABW81864 Tumour-as
45	6806	99.9	1255	9 ADW28640	ADW28640 HER2, SEQ

## ALIGNMENTS

RESULT 1	AAW01111	standard, protein, 1255 AA.
ID	AAW01111	
AC	AAW01111	
XX		
DT	01-JAN-1997	(first entry)
XX		
DE	HER-2/neu protein.	
XX		
KM	HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;	
KW	breast cancer; ovary cancer; colon cancer; lung cancer; prostate cancer;	
XX	immunisation; tumour; vaccine; vector.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Domain	676..1255
FT		/label= "intracellular domain
FT		/note= "claimed domain, useful for immunisation"
XX		
PN	WO9630514-A1.	
XX		
PD	03-OCT-1996.	
XX		
PF	28-MAR-1996;	96WO-US001689.
XX		
PR	31-MAR-1995;	95US-00414417.
XX		
PA	(UNITW ) UNIV WASHINGTON.	
XX		
PI	Cheever MA, Disis ML;	
XX		
DR	WPI: 1996-455361/45.	
XX		
DR	N-PSDB; AAT40739.	
XX		
PT	DNA encoding HER-2/neu poly:peptide(s) - used for prevention or treatment	
PT	of malignancies with which the HER-2/neu oncogene is associated.	
XX		
PS	Claim 2: Page 56-61; 71pp; English.	
XX		
CC	Human HER-2/neu protein (AAW01111), also called p185, or c-erbB2, is the	
CC	product of the HER-2/neu oncogene (see also AAT40739). The protein is	
CC	over-expressed in various cancers, including breast, ovarian, colon, lung	
CC	and prostate. The intracellular domain of the protein can be used to	
CC	immunise an animal against a malignancy with which the oncogene is	
CC	associated. The polypeptide can be produced in transformed host cells for	

CC use in immunisation. Alternatively, animal cells are transfected in vivo  
 CC or ex vivo with a viral vector that directs expression of the polypeptide  
 xx  
 xx Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 2; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELALCRWGLLALLPQASTOVCTGDMKLRIPASPEHTLDMLRHLYGCGVQVQNL 60
DB 1 MELALCRWGLLALLPQASTOVCTGDMKLRIPASPEHTLDMLRHLYGCGVQVQNL 60
QY 61 ELTYLPTNASLFLDIOEVGVVLIANQVQVPLQRLIVRGTLFEDNYALAVLNG 120
DB 61 ELTYLPTNASLFLDIOEVGVVLIANQVQVPLQRLIVRGTLFEDNYALAVLNG 120
QY 121 DPLNNTPTVTGASPGGLRELQLRLSLTEILKGGVLIQRNPQLCYODTIIMKDI FHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQLRLSLTEILKGGVLIQRNPQLCYODTIIMKDI FHKNNOLA 180
QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCOSLTRVCAGGACRCKGPLEPTDCHEQC 240
DB 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCOSLTRVCAGGACRCKGPLEPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACP 300
QY 301 YNYLSTDVSGCTLVCPRLHNOEVTABDGTQRCCKSPCARVCYGLGMEHLREVRATVSAN 360
DB 301 YNYLSTDVSGCTLVCPRLHNOEVTABDGTQRCCKSPCARVCYGLGMEHLREVRATVSAN 360
QY 361 IOEFAGCKKIFESLAFIPSPFDGDPASNTAPLOPQLOVFETLEITGLYLSAPDLSLP 420
DB 361 IOEFAGCKKIFESLAFIPSPFDGDPASNTAPLOPQLOVFETLEITGLYLSAPDLSLP 420
QY 421 DLSTVQNLQVIRGRILHNGAVSLTQGLISWLGRLSRLSGSLALIHNTHLCFVHTV 480
DB 421 DLSTVQNLQVIRGRILHNGAVSLTQGLISWLGRLSRLSGSLALIHNTHLCFVHTV 480
QY 481 PWDOLFRNPHQALHTANRBEDECVBEGIAHQLCARGHCMWPGPTQCNCSQPLRGQBC 540
DB 481 PWDOLFRNPHQALHTANRBEDECVBEGIAHQLCARGHCMWPGPTQCNCSQPLRGQBC 540
QY 541 VEECEVLQGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCVACAHYKDPFCVAKC 600
DB 541 VEECEVLQGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCVACAHYKDPFCVAKC 600
QY 601 PSGVPRDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDDKGPABQORASPLTSLISAVVG 660
DB 601 PSGVPRDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDDKGPABQORASPLTSLISAVVG 660
QY 661 ILVVVLGVVFGIILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOQMRILKETEL 720
DB 661 ILVVVLGVVFGIILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOQMRILKETEL 720
QY 721 RKVKTLSGSGAGTYVKKGIPIRGENVKIPVAIKVIRENTSPRANKELIDEAVVMGVGSP 780
DB 721 RKVKTLSGSGAGTYVKKGIPIRGENVKIPVAIKVIRENTSPRANKELIDEAVVMGVGSP 780
QY 781 YVSRLLGICLTSTVOLVTQLMPYGCILDHVRNRRGLSGODLLNMCQIAKMSYLEDVR 840
DB 781 YVSRLLGICLTSTVOLVTQLMPYGCILDHVRNRRGLSGODLLNMCQIAKMSYLEDVR 840
QY 841 LVHRDLAARNVLVKSPPNHVKTIDFGLARLLIDETBYTAADGKVPKIMWALESLIRRF 900
DB 841 LVHRDLAARNVLVKSPPNHVKTIDFGLARLLIDETBYTAADGKVPKIMWALESLIRRF 900
QY 901 HOSDWSYGVTVWELMTFGAKPYDGI PAEIPDLLEKGRLLPQPICTIDVMIWVKCM 960
DB 901 HOSDWSYGVTVWELMTFGAKPYDGI PAEIPDLLEKGRLLPQPICTIDVMIWVKCM 960

```

```

QY 961 IDSECRPRFRELVEFSRMAADPQRFVVIQHEDIGPASPUDSTYRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRFRELVEFSRMAADPQRFVVIQHEDIGPASPUDSTYRSLLEDDMDGLVDA 1020
QY 1021 EBYLVPOQGFPCPPAPACGMVHRRSSSTRSGGDLTGLRPSSEEARSPAPASEG 1080
DB 1021 EBYLVPOQGFPCPPAPACGMVHRRSSSTRSGGDLTGLRPSSEEARSPAPASEG 1080
QY 1081 AGSDVFDLDMGAAGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPOPEYV 1140
DB 1081 AGSDVFDLDMGAAGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPOPEYV 1140
QY 1141 NQPDVPRQPPSPRRCPLPAPAPAGATLERPKTSLPGKXGVKDYAPFGGAVENBEYLTPO 1200
DB 1141 NQPDVPRQPPSPRRCPLPAPAPAGATLERPKTSLPGKXGVKDYAPFGGAVENBEYLTPO 1200
QY 1201 GGAAPQHPHPAPFAPADNLYYWDODPPERAPASTPGTGTANPEYLGIDVPV 1255
DB 1201 GGAAPQHPHPAPFAPADNLYYWDODPPERAPASTPGTGTANPEYLGIDVPV 1255

RESULT 2
AAW92406
ID AAW92406 standard; protein; 1255 AA.
XX
AC AAW92406;
XX
DT 21-APR-1999 (first entry)
XX
DE Human HER-2/neu oncogene protein.
XX
KW HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
KW malignancy; treatment; tumour.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 676..1255
FT /note="region which elicits immune response"
XX
PN US5869445-A.
XX
PD 09-FEB-1999.
XX
PE 01-APR-1996; 96US-00625101.
XX
PR 17-MAR-1993; 93US-00033644.
PR 12-AUG-1993; 93US-00106112.
PR 31-MAR-1995; 95US-00414417.
XX
PA (UNIV ) UNIV WASHINGTON.
XX
PI Cheever MA, Disis ML;
XX
DR WPI; 1999-152835/13.
DR N-PSDB; AAX01912.
XX
PT Use of HER-2/neu polypeptides - for eliciting an immune response to an
PT HER-2/neu associated malignancy, particularly for treating or preventing
PT tumour.
XX
PS Claim 3; Col 31-38; 26pp; English.
XX
CC This sequence represents the human HER-2/neu oncogene protein. A fragment
CC of this protein is used in a method for eliciting or enhancing an immune
CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
CC B cells to produce an immune response to the HER-2/neu protein. The
CC method can be used for immunisation against a malignancy in which the HER
CC -2/neu oncogene is associated and in the treatment of an existing tumour,
CC or to prevent tumour occurrence or recurrence
XX
SQ Sequence 1255 AA;

```

Query Match	100.0%;	Score 6815;	DB 2;	Length 1255;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1255;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MEIAAI	CRWGLL	LALLP
DB	1	MEIAAI	CRWGLL	LALLP
QY	61	ELTYLP	PTNAST	SFLDOI
DB	61	ELTYLP	PTNAST	SFLDOI
QY	121	DLNNTT	PTYGAS	PGGLRE
DB	121	DLNNTT	PTYGAS	PGGLRE
QY	181	LTLIDT	NRSRAC	PCSPK
DB	181	LTLIDT	NRSRAC	PCSPK
QY	241	AAGCTP	PKHSDC	LACI
DB	241	AAGCTP	PKHSDC	LACI
QY	301	YNYLST	DVGSCT	LCPLH
DB	301	YNYLST	DVGSCT	LCPLH
QY	361	IOEPAG	CKKIFG	SLAE
DB	361	IOEPAG	CKKIFG	SLAE
QY	421	DLSPON	LOVIRG	RII
DB	421	DLSPON	LOVIRG	RII
QY	481	PMDOLE	FRNHOL	LHTANR
DB	481	PMDOLE	FRNHOL	LHTANR
QY	541	VEECRV	LOGLPRE	YVNA
DB	541	VEECRV	LOGLPRE	YVNA
QY	601	PEGVKE	DLSPYM	PIWKPE
DB	601	PEGVKE	DLSPYM	PIWKPE
QY	661	ILLVVL	GVVFGI	LIRK
DB	661	ILLVVL	GVVFGI	LIRK
QY	721	RKVKVL	SGSAGF	VTYKGI
DB	721	RKVKVL	SGSAGF	VTYKGI
QY	781	YYSRLI	GICLTST	VO
DB	781	YYSRLI	GICLTST	VO
QY	841	LVRHDA	ANRV	LVKS
DB	841	LVRHDA	ANRV	LVKS
QY	901	HOSDVS	YGVTVL	MTFG
DB	901	HOSDVS	YGVTVL	MTFG
QY	961	IDSECP	RRREL	VSERS
DB	961	IDSECP	RRREL	VSERS
QY	1021	BEVLVP	QOGF	PCPD
DB	1021	BEVLVP	QOGF	PCPD

Db	1021	BEYLVPOQGFPCDPAPAGAGMWHRRRSSSTRSGGDLTLGLEPSEBEAPRSILAPSEG	1080
Qy	1081	AGSDVPDDGLGMGAKKGLQSLPTHDSPLQRYSEDPTVPLPSEYDGYVAPLTCSPQBEYV	1140
Db	1081	AGSDVPDDGLGMGAKKGLQSLPTHDSPLQRYSEDPTVPLPSEYDGYVAPLTCSPQBEYV	1140
Qy	1141	NOVDVRPOPSPREGRPLPAARPAATLERRPKTLLSGKNGVVDVFAFGAIVENPEYLTPO	1200
Db	1141	NOVDVRPOPSPREGRPLPAARPAATLERRPKTLLSGKNGVVDVFAFGAIVENPEYLTPO	1200
Qy	1201	GGAAPQHPHPAPSPAFDNLYYMQDDPERGAPSTFGKTPTAENPEYLGIDVPU	1255
Db	1201	GGAAPQHPHPAPSPAFDNLYYMQDDPERGAPSTFGKTPTAENPEYLGIDVPU	1255

RESULT 3  
AA584780  
ID AA584780 standard; protein; 1255 AA.

XX  
AC AA584780;

DT 08-AUG-2000 (first entry)

XX Amino acid sequence of the SPLICE erdb-2 receptor protein.

DE SPLICE erdb-2 receptor protein; cell transformation disorder; cancer;

KW tumor cell proliferation; tissue degeneration; arthropathy;

KW bone resorption; inflammatory disease; degenerative disorder;

XX wound healing.

OS Homo sapiens.

XX MO200020579-A1.

XX 13-APR-2000.

XX 01-OCT-1999; 99WC-CA000912.

XX 02-OCT-1998; 98US-00165192.

XX (UYMC-) UNIV MCMASTER.

XX PA Muller WJ, Siegel PM;

XX PI WPI, 2000-303768/26.

XX DR N-PSDB; AAA14812.

XX PT Nucleic acid encoding an erdb 2 receptor protein designated SPLICE erdb-

XX PT 2, inhibitors of the protein are useful for treatment of cancer.

XX PS Claim 3; Fig 2; 60pp; English.

XX The present sequence represents a SPLICE erdb-2 receptor protein. The  
XX protein has an in-frame deletion of 16 amino acids, 2 of which are  
XX conserved cysteine residues, compared to the unspliced protein. The erdb-  
XX 2 polynucleotide is used to construct probes for detecting disorders of  
XX cell transformation such as cancer. Antibodies to the protein may be used  
XX to detect SPLICE erdb-2 in a sample. Agents (e.g. antisense  
XX oligonucleotides) which inhibit the expression of SPLICE erdb-2 are  
XX useful for reducing tumor cell proliferation and treating cancer.  
XX Substances which stimulate SPLICE erdb-2 are useful for treating  
XX conditions involving damaged cells including conditions in which  
XX degeneration of tissue occurs, such as arthropathy, bone resorption,  
XX inflammatory diseases, degenerative disorders of the central nervous  
XX system and wound healing

XX Sequence 1255 AA:

Query Match 100.0%; Score 6815; DB 3; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy	1	MELAAICRMGLIALALLP	PGAASTOVCTGTDMKRLRLPASPEHILDMLEKHLVQCGVQVGNL	60
Db	1	MELAAICRMGLIALALLP	PGAASTOVCTGTDMKRLRLPASPEHILDMLEKHLVQCGVQVGNL	60
Qy	61	ELTYLPTNASTL	SFLODIOEVQGYVLIHANNVOYVPLRLRLVRGSTOLFEDNYALAVLDNG	120
Db	61	ELTYLPTNASTL	SFLODIOEVQGYVLIHANNVOYVPLRLRLVRGSTOLFEDNYALAVLDNG	120
Qy	121	DPLANTTPVTGAS	PGGLRELOLSLTELKGGVLIQBNPOLCYODITLMKDIFHKNQOLA	180
Db	121	DPLANTTPVTGAS	PGGLRELOLSLTELKGGVLIQBNPOLCYODITLMKDIFHKNQOLA	180
Qy	181	LTLIDTNRSPACHPC	SPMCKGSRWCGSSSEDQSLTFTVCAGGACARCKGRLPTCCHEQC	240
Db	181	LTLIDTNRSPACHPC	SPMCKGSRWCGSSSEDQSLTFTVCAGGACARCKGRLPTCCHEQC	240
Qy	241	AAGCTBPKSDCLAC	LHFHNSGICELHCPALVTYNTDTPESMPNPEGRTFTGASCVAACP	300
Db	241	AAGCTBPKSDCLAC	LHFHNSGICELHCPALVTYNTDTPESMPNPEGRTFTGASCVAACP	300
Qy	301	YNYLSTDVSGCTL	VCPLHNOEVTAEADGTORCEKSKCARCYAGLGMHLEREAVTSAN	360
Db	301	YNYLSTDVSGCTL	VCPLHNOEVTAEADGTORCEKSKCARCYAGLGMHLEREAVTSAN	360
Qy	361	IOEFACCKKIFGSLA	FLPESFDGDPASNTVPLQEOULQVEFTLEITGYLTSAMPDSL	420
Db	361	IOEFACCKKIFGSLA	FLPESFDGDPASNTVPLQEOULQVEFTLEITGYLTSAMPDSL	420
Qy	421	DLSVPNLOVIRRI	LHNGAYSLTLQGLGSMGLRLRELGSGLALIHNTHLCPNHTV	480
Db	421	DLSVPNLOVIRRI	LHNGAYSLTLQGLGSMGLRLRELGSGLALIHNTHLCPNHTV	480
Qy	481	PMQDLFRNPHQAL	LHTANRPEDECVGEGLACHOLCARGHGWPPTOCVNCQPLRGEC	540
Db	481	PMQDLFRNPHQAL	LHTANRPEDECVGEGLACHOLCARGHGWPPTOCVNCQPLRGEC	540
Qy	541	VEBCRYLQGLPRE	VYNAHCLPCHBECQPONGSYTCGPBEADQCAAHYKDPFCVARC	600
Db	541	VEBCRYLQGLPRE	VYNAHCLPCHBECQPONGSYTCGPBEADQCAAHYKDPFCVARC	600
Qy	601	PSGVKRDLSYMP	IMKPRDEEGAOCPCINCHSCVDLDNGCPAEOGASPLTSLISAVNG	660
Db	601	PSGVKRDLSYMP	IMKPRDEEGAOCPCINCHSCVDLDNGCPAEOGASPLTSLISAVNG	660
Qy	661	ILLVNVLVGVFGLI	IKRROOKIRKYTYRRLLOETELVEPLTPSGAMPNOAMRLIKETEL	720
Db	661	ILLVNVLVGVFGLI	IKRROOKIRKYTYRRLLOETELVEPLTPSGAMPNOAMRLIKETEL	720
Qy	721	RKYKVLGSGAF	GVTVYKGIWIPDGENVYKIPVALIKVLRENTSPKANKEILDEAYVAGVSP	780
Db	721	RKYKVLGSGAF	GVTVYKGIWIPDGENVYKIPVALIKVLRENTSPKANKEILDEAYVAGVSP	780
Qy	781	YYSRLILGICL	TSIVOLVTOCLMPFGCLLDHYRENGRIGSQDILMWCMQIANGMSYLEDVR	840
Db	781	YYSRLILGICL	TSIVOLVTOCLMPFGCLLDHYRENGRIGSQDILMWCMQIANGMSYLEDVR	840
Qy	841	LHYRDLAARVNL	VKSPMHVKITTPGRLARLDDIDETEHADGKVPIMMALESILRRFT	900
Db	841	LHYRDLAARVNL	VKSPMHVKITTPGRLARLDDIDETEHADGKVPIMMALESILRRFT	900
Qy	901	HOSDWSYGVTV	WELMTFGAKPYDGI	960
Db	901	HOSDWSYGVTV	WELMTFGAKPYDGI	960
Qy	961	IOSECBPRFREL	VEFSRMAADPORFVYIQNEDLGPASPLDSTYRSLLEBDDMDGLVDA	1020
Db	961	IOSECBPRFREL	VEFSRMAADPORFVYIQNEDLGPASPLDSTYRSLLEBDDMDGLVDA	1020
Qy	1021	BEVLVYQOQF	FCDDPAPAGAMVHHRRSSSTRSGGDLTGLEPSEBEARSPILASBEG	1080
Db	1021	BEVLVYQOQF	FCDDPAPAGAMVHHRRSSSTRSGGDLTGLEPSEBEARSPILASBEG	1080
Qy	1081	AGSDVDPDGL	GMGAJKGLPTHDSPLOQYSEDPVPLPSEFSDGVYAPLTCSPQEPYV	1140
Db	1081	AGSDVDPDGL	GMGAJKGLPTHDSPLOQYSEDPVPLPSEFSDGVYAPLTCSPQEPYV	1140

Db	1081	AGSDVFDGDLGAAKGLQSLFTHDPSLPQRTSEDFYFVPSFTDGYAPLTCSPQPEYV	114
Qy	1141	NQPDVRPQPPEPSREBPPLPAARFAGATLERPKTLSPGKGVWDVPAFGAVENPEYLTPQ	120
Db	1141	NQPDVRPQPPEPSREBPPLPAARFAGATLERPKTLSPGKGVWDVPAFGAVENPEYLTPQ	120
Qy	1201	GGAAPQHPHPAPFASAFDNLVYWDODPPBRGAPPESTFKCTPTAENPEYLGLDVVP	1255
Db	1201	GGAAPQHPHPAPFASAFDNLVYWDODPPBRGAPPESTFKCTPTAENPEYLGLDVVP	1255
RESULT 4			
ID	AAAB21198	standard; protein; 1255 AA.	
AC	AAAB21198;		
DT	12-JAN-2001	(first entry)	
DE	Human HER-2/neu protein.		
KM	Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;		
KW	breast cancer; prostate cancer; ovarian cancer; lung cancer;		
KM	colon cancer.		
OS	Homo sapiens.		
PN	WO200044899-A1.		
XX	03-AUG-2000.		
PF	28-JAN-2000; 2000WO-US002164.		
PR	29-JAN-1999; 99US-0117976P.		
XX	(CORI-) CORIXA CORP.		
PA	(SMK ) SMITHKLINE BEECHAM.		
PI	Cheever MA, Gheysen D;		
XX	WPI. 2000-505976/45.		
DR	N-PSDB; AAA89736.		
XX	HER-2/neu extracellular domain/phosphorylation domain fusion proteins		
PT	useful for vaccinating against breast, ovarian, colon, lung and prostate		
PT	cancers.		
XX	Claim 52; Fig 7; 128pp; English.		
PS	The present sequence is the human HER-2/neu protein. It is a member of		
XX	the tyrosine kinase family of receptor-like glycoproteins and shows		
CC	homology to the epidermal growth factor receptor (EGFR). It probably		
CC	plays a part in cell growth and/or differentiation. The HER-2/neu gene is		
CC	an oncogene. An HER-2/neu fusion protein comprising a HER-2/neu		
CC	extracellular domain fused to a HER-2/neu phosphorylation domain may be		
CC	used to treat or prevent cancer by eliciting or enhancing an immune		
CC	response to the HER-2/neu protein. It may be used to treat malignancies		
CC	such as breast, ovarian, colon, lung and prostate cancers, and may be		
XX	used as an antigen to vaccinate against these neoplasias		
XX	Sequence 1255 AA;		
Qy	Query Match	100.0%; Score 6815; DB 3; Length 1255;	
Db	Best Local Similarity	100.0%; Pred. No. 0; Indels 0; Gaps 0	
Db	Matches 1255; Conservative	0; Mismatches 0; Indels 0; Gaps 0	
Qy	1	MELALACRWGLLIALPGCASTVCTGTDMLRLPASBETHLDMRLHYOGGVQGNL	60
Db	1	MELALACRWGLLIALPGCASTVCTGTDMLRLPASBETHLDMRLHYOGGVQGNL	60
Qy	61	ELTVPYTNASLSFLDIDQEVGVLIHANOVRPLQRLRIYRGTVLFEDNTALAVLDNG	120

```

Db      61 ELTYLPTNASLSFLQDIOEVGVVLIHNVQVPLQRLIRVRGIQLFEDNYALAVLDNG 120
Qy      121 DDLNNTTPTTGASPGSLRELQRLSLTEILKGVLLIORNPOLCYODITLIMKDI FHKNNQIA 180
Db      121 DDLNNTTPTTGASPGSLRELQRLSLTEILKGVLLIORNPOLCYODITLIMKDI FHKNNQIA 180
Qy      181 LTLIDTNRBRACHPCSPMKGSRCMGSESEDCCSLTRTVCAAGCARCKGPLPTDCCHEQC 240
Db      181 LTLIDTNRBRACHPCSPMKGSRCMGSESEDCCSLTRTVCAAGCARCKGPLPTDCCHEQC 240
Qy      241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTESMPNPEGRYTFGASCVTACP 300
Db      241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTESMPNPEGRYTFGASCVTACP 300
Qy      301 YNYLSTDVGSCTLVNCPHNOEVTABDGTORCKSPCARVCYGLGMEHLREVRATSN 360
Db      301 YNYLSTDVGSCTLVNCPHNOEVTABDGTORCKSPCARVCYGLGMEHLREVRATSN 360
Qy      361 IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLOPELOVFEITLGYLISAMPDSL 420
Db      361 IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLOPELOVFEITLGYLISAMPDSL 420
Qy      421 DLSVFQNLQVIRGRILHNGAVSLTLOGLISWLGRLSLRELGSGLALIHNTHLCEVHTV 480
Db      421 DLSVFQNLQVIRGRILHNGAVSLTLOGLISWLGRLSLRELGSGLALIHNTHLCEVHTV 480
Qy      481 PMDQLFRRPHOALLHTANRPEDECEGEGLAGHQLCARHGCMNGRPPQCNCQPLRGDGC 540
Db      481 PMDQLFRRPHOALLHTANRPEDECEGEGLAGHQLCARHGCMNGRPPQCNCQPLRGDGC 540
Qy      541 VEECEVTLQGLPREVYNARHCLPCHPECPONGSVTCFGEALQCVCAHYKDPPEVCVARC 600
Db      541 VEECEVTLQGLPREVYNARHCLPCHPECPONGSVTCFGEALQCVCAHYKDPPEVCVARC 600
Qy      601 PSGVPRDLSEYMPIMKPEPDEEGACQPCPINCTHSVDLDDKGCPEAGRASPLTISIISAVVG 660
Db      601 PSGVPRDLSEYMPIMKPEPDEEGACQPCPINCTHSVDLDDKGCPEAGRASPLTISIISAVVG 660
Qy      661 ILLVVLGVVREGILLTKRQOKIRKTYMRLLQETELVEPLTSGAMPNOAQRILKETEL 720
Db      661 ILLVVLGVVREGILLTKRQOKIRKTYMRLLQETELVEPLTSGAMPNOAQRILKETEL 720
Qy      721 RKVKYLSGSAFGTVYKGIWIPDGENYKIPVAIKVIRENTSPRANKELIDEAYVMAGVSP 780
Db      721 RKVKYLSGSAFGTVYKGIWIPDGENYKIPVAIKVIRENTSPRANKELIDEAYVMAGVSP 780
Qy      781 YVSRLLGICLTSTVQLVTQLMFYGCLLDHVRENRGLSQDILLNMCQIAKGMSTYLEDVR 840
Db      781 YVSRLLGICLTSTVQLVTQLMFYGCLLDHVRENRGLSQDILLNMCQIAKGMSTYLEDVR 840
Qy      841 LVHRDLAARNVTVKSGPNHVKITDFGLARLLIDETETAHADGKVPFKMMALSSILARRPT 900
Db      841 LVHRDLAARNVTVKSGPNHVKITDFGLARLLIDETETAHADGKVPFKMMALSSILARRPT 900
Qy      901 HOSDWSYGVTVWELMTFGAKXPDDGIPAREIPDLLEKGERLPOPPICITDVMIMWIKCM 960
Db      901 HOSDWSYGVTVWELMTFGAKXPDDGIPAREIPDLLEKGERLPOPPICITDVMIMWIKCM 960
Qy      961 IDSECRPRELVSSESRMARDPORFVVIQNEIDLGPASPLDSTFYRSLLEDDMGDLVDA 1020
Db      961 IDSECRPRELVSSESRMARDPORFVVIQNEIDLGPASPLDSTFYRSLLEDDMGDLVDA 1020
Qy      1021 EBYLVPOQGFPCPDPAFGAGVNHHRSSSTRSGGGLTLGLPESEEAAPSPPLAPSG 1080
Db      1021 EBYLVPOQGFPCPDPAFGAGVNHHRSSSTRSGGGLTLGLPESEEAAPSPPLAPSG 1080
Qy      1081 AGSDVFDGSLGMAAGLQSLPHTDPSPLQRYSEDPTVLPSTETDGYAPLTCSPPEYV 1140
Db      1081 AGSDVFDGSLGMAAGLQSLPHTDPSPLQRYSEDPTVLPSTETDGYAPLTCSPPEYV 1140
Qy      1141 NOPDVAPQPSPREGFLPAARPAAGATLERPKTILSPGKGVKVDVAFGAVENPEYLQ 1200
Db      1141 NOPDVAPQPSPREGFLPAARPAAGATLERPKTILSPGKGVKVDVAFGAVENPEYLQ 1200

```

```

Qy      1201 GGAAPQHPPPAFSPAFDNLYYMDODPPERGAPSTFKGTFRANPEYLGLDVPV 1255
Db      1201 GGAAPQHPPPAFSPAFDNLYYMDODPPERGAPSTFKGTFRANPEYLGLDVPV 1255

RESULT 5
AAG88267
ID AAG88267 standard; protein; 1255 AA.
XX
AC AAG88267;
XX
DT 11-SEP-2001 (first entry)
XX
DE HER2/neu amino acid sequence.
XX
KW Human; HER2/neu; epitope; human leukocyte antigen; HLA; T cell;
KW immune response; vaccine; tumour; cancer; cytostatic; immunostimulant;
KW tumour-associated antigen; T lymphocyte; cytotoxic T lymphocyte; CTL.
XX
OS Homo sapiens.
XX
PN WO200141787-A1.
XX
PD 14-JUN-2001.
XX
PF 11-DEC-2000; 2000WO-US033591.
XX
PR 10-DEC-1999; 99US-00458299.
XX
PA (EPRIM-) EPRIMUNE INC.
XX
PI Fikes J, Sette A, Sidney J, Southwood S, Chesnut R, Celis E;
PI Keogh E;
XX
DR WPI: 2001-374995/39.
XX
PT An isolated prepared HER2/neu epitope useful in a vaccine for inducing
PS cellular immune responses for the prevention and treatment of cancer.
XX
PS Disclosure; Page 15; 1999p; English.
XX
CC The present invention describes isolated prepared HER2/neu epitopes (I).
CC Also described are: (1) a clonal cytotoxic T lymphocyte (CTL) that is
CC culture in vitro and binds to a complex of an epitope (II), bound to a
CC human leukocyte antigen (HLA) molecule; (2) a peptide (II) comprising (I)
CC and a second epitope and the peptide is less than 50 contiguous amino
CC acids that have 100% identity with a native peptide sequence of HER2/neu;
CC (3) a vaccine composition (III) comprising (II) and a pharmaceutical
CC excipient; (4) an isolated nucleic acid encoding a peptide comprising (I)
CC and (5) an isolated nucleic acid encoding (II). (I) has cytostatic and
CC immunostimulant activities, and can be used in vaccines. (II), (II) and
CC (III) are useful for inducing cellular immune responses for the
CC prevention and treatment of cancer. (I) and (II) are useful for
CC monitoring or evaluating an immune response to a tumour-associated
CC antigen when incubated with a T lymphocyte sample form a patient and
CC detecting the presence of bound T lymphocyte to (I) or (II). Epitope
CC based vaccines mean that immunosuppressive epitopes that may be present
CC in whole antigens may be avoided. Selected epitopes may be combined to
CC enhance immunogenicity. The possible pathological side effects caused by
CC infectious agents or whole protein antigen is eliminated. The vaccine
CC provides the ability to direct and focus an immune response to multiple
CC selected antigens from the same pathogen. Epitope-based anti-tumour
CC vaccines provides the opportunity to combine epitopes derived from
CC multiple tumour-associated molecules addressing the problem of tumour-
CC tumour variability and reducing the likelihood of tumour escape due to
CC antigen loss. AAG88266 to AAG89121 represent amino acid sequences used in
CC the exemplification of the present invention
XX
SQ Sequence 1255 AA;

```

```

Query Match      100.0%; Score 6815; DB 4; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;

```



QY 121 DPLNNTTEVTGASPGGLRELAQLRSLTEILKGGVLIQBNPOLCYOPTILMKDIFHKNNOLA 180  
 DB 121 DPLNNTTEVTGASPGGLRELAQLRSLTEILKGGVLIQBNPOLCYOPTILMKDIFHKNNOLA 180  
 QY 181 LTLIDTNSRACHPCS PMCKGSRGWSESESDCOSLTRTVCAAGCARCKGPLPTDCHEQC 240  
 DB 181 LTLIDTNSRACHPCS PMCKGSRGWSESESDCOSLTRTVCAAGCARCKGPLPTDCHEQC 240  
 QY 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTPESNPNEGRTTFCASCTYACP 300  
 DB 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTPESNPNEGRTTFCASCTYACP 300  
 QY 301 YNYLSTDVSGCTLVCPLNHOEVTAEADGTORCEKCKPCARVCYCGMEHLREVRVATSAN 360  
 DB 301 YNYLSTDVSGCTLVCPLNHOEVTAEADGTORCEKCKPCARVCYCGMEHLREVRVATSAN 360  
 QY 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPELOVFTLEBITGYLISAMPDSL 420  
 DB 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPELOVFTLEBITGYLISAMPDSL 420  
 QY 421 DLSVQNLQVTRGRILHNGAYSLTQIGISWLGRLSRLBELSGSLALIHNTHLCPVHTV 480  
 DB 421 DLSVQNLQVTRGRILHNGAYSLTQIGISWLGRLSRLBELSGSLALIHNTHLCPVHTV 480  
 QY 481 PMDOLFRPHOALHTANRPEDECVGEGLACHQLCARGHCMGPGPTOCVNCGQPLRGQEC 540  
 DB 481 PMDOLFRPHOALHTANRPEDECVGEGLACHQLCARGHCMGPGPTOCVNCGQPLRGQEC 540  
 QY 541 VBECEVLOGLPREYVNAHCLPCHPECOQNGSVTCFGEADQCVACAHYKDPPECVAR 600  
 DB 541 VBECEVLOGLPREYVNAHCLPCHPECOQNGSVTCFGEADQCVACAHYKDPPECVAR 600  
 QY 601 PSGVPRDLSYMPIMKFPBEBGACQPCPINTCHSCVDLDDKCPAQRASPLTSTISAVG 660  
 DB 601 PSGVPRDLSYMPIMKFPBEBGACQPCPINTCHSCVDLDDKCPAQRASPLTSTISAVG 660  
 QY 661 ILLVVVLGVVGGILIKRBOQKIRKTYMRLLQETLVEPLPSSGAMPNOAQRILKEKTEL 720  
 DB 661 ILLVVVLGVVGGILIKRBOQKIRKTYMRLLQETLVEPLPSSGAMPNOAQRILKEKTEL 720  
 QY 721 RKVKVLSGAGFTVYKGIWIPDGENVKIPVAIKVLRNTSPYANKELIDEAVMAGVSP 780  
 DB 721 RKVKVLSGAGFTVYKGIWIPDGENVKIPVAIKVLRNTSPYANKELIDEAVMAGVSP 780  
 QY 781 YVSRLLGICLTSTVOLVQLMPYGCCLDHVRENRLGSSODILANCMQIAKMSYLBVR 840  
 DB 781 YVSRLLGICLTSTVOLVQLMPYGCCLDHVRENRLGSSODILANCMQIAKMSYLBVR 840  
 QY 841 LVHRLARBNVLVKS PNHVKITDPLGLARLDDIDETRYHADGSKVITKMMALSTILRRFT 900  
 DB 841 LVHRLARBNVLVKS PNHVKITDPLGLARLDDIDETRYHADGSKVITKMMALSTILRRFT 900  
 QY 901 HQSDVMSGVTVWELMTFGAKPYDGI PAREIPDLLEKESRLPQPICTIDVYIMVKKCM 960  
 DB 901 HQSDVMSGVTVWELMTFGAKPYDGI PAREIPDLLEKESRLPQPICTIDVYIMVKKCM 960  
 QY 961 IDSECRPPRELIVSFSRMAARDPQRFVYIQNEIDLCPASPLDSTFYRSLLEDDMDGLVDA 1020  
 DB 961 IDSECRPPRELIVSFSRMAARDPQRFVYIQNEIDLCPASPLDSTFYRSLLEDDMDGLVDA 1020  
 QY 1021 EBYLVPOQGFPCPDPAAGAGVHHRHSSSTRSGGDLTLGLSESEBAPRSLAPBEG 1080  
 DB 1021 EBYLVPOQGFPCPDPAAGAGVHHRHSSSTRSGGDLTLGLSESEBAPRSLAPBEG 1080  
 QY 1081 AGSDVPDDDLGWAAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
 DB 1081 AGSDVPDDDLGWAAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
 QY 1141 NOPDVRPOPSPREBGPLPAAPAGATLERPKTLSFGKGVKQVQVAPFGAGAVENPEYLTPO 1200  
 DB 1141 NOPDVRPOPSPREBGPLPAAPAGATLERPKTLSFGKGVKQVQVAPFGAGAVENPEYLTPO 1200

QY 1201 GGAAPQHPHPAPFAFPAFNLVYMDPPERGAPESTFGKTPAENPEYLGIDVPV 1255  
 DB 1201 GGAAPQHPHPAPFAFPAFNLVYMDPPERGAPESTFGKTPAENPEYLGIDVPV 1255  
 RESULT 7  
 AAE20479  
 ID AAE20479 standard; protein; 1255 AA.  
 AC AAE20479;  
 XX  
 DT 01-JUL-2002 (first entry)  
 XX  
 DE Human Her-2/neu protein.  
 XX  
 KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
 XX human leukocyte antigen; HLA; vaccine; malignancy; cytotoxic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1021..1030  
 FT /note="Naturally processed HLA-B44-restricted epitope"  
 PN WO200214503-A2.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 14-AUG-2001; 2001WO-US041733.  
 XX  
 PR 14-AUG-2000; 2000US-0225152P.  
 PR 28-SEP-2000; 2000US-0236428P.  
 PR 21-FEB-2001; 2001US-0270520P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PI Hand-Zimmermann S, Cheever MA, Foy TM, Lodes MD, Kalos MD;  
 PI McNeill PD, Veddyck TS;  
 PI WPI; 2002-280756/32.  
 DR N-PSDB; AAD32743.  
 XX  
 PS Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
 PT prevention and diagnosis of cancer, preferably breast cancer.  
 PT  
 XX  
 PS Disclosure: Page 114-117; 129pp; English.  
 CC The invention relates to an isolated Her-2/Neu polypeptide composition  
 CC effective for eliciting an immune response. The invention is useful for  
 CC eliciting an immune response in a patient, where the patient is human  
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
 CC The composition is useful for the therapy and diagnosis of cancer,  
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
 CC and other compositions, for stimulating and/or expanding T cells specific for  
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
 CC patient. The invention is useful for stimulating a T cell response in a  
 CC human patient, as probe or primer for nucleic acid hybridisation, to  
 CC selectively form duplex molecules with complementary stretches of the  
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
 CC length gene from a suitable library, and to direct expression of a  
 CC polypeptide in appropriate host cells. The composition is useful in  
 CC prophylactic or therapeutic applications and for the treatment of cancer,  
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
 CC associated malignancies. The invention is useful in gene therapy. The  
 CC present sequence is human Her-2/neu protein  
 XX  
 SO Sequence 1255 AA:  
 Query Match 100.0%; Score 6815; DB 5; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELAALCRWGILLALLLPGAASVQVCTGTDMKRLPASPETHLDMRLRLYOGCQVVGNTL 60
Db 1 MELAALCRWGILLALLLPGAASVQVCTGTDMKRLPASPETHLDMRLRLYOGCQVVGNTL 60
QY 61 ELTYLPTNASISFLDIOEVQGYVLIANHVRQVPLQRLRIVRGQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASISFLDIOEVQGYVLIANHVRQVPLQRLRIVRGQLFEDNYALAVLDNG 120
QY 121 DPLNNTTAVTGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODPTLMKDIFFHKNNOLA 180
Db 121 DPLNNTTAVTGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODPTLMKDIFFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPCKSGSRCWSSSEDCQSLTRTVACGAGCARCKGRLPTDCHEQC 240
Db 181 LTLIDTNRSRACHPCSPCKSGSRCWSSSEDCQSLTRTVACGAGCARCKGRLPTDCHEQC 240
QY 241 AAGCTGPRHSDCLAFHNSGICELHCPALVTYNTDFFESMPNBEGRYTGASCVTACP 300
Db 241 AAGCTGPRHSDCLAFHNSGICELHCPALVTYNTDFFESMPNBEGRYTGASCVTACP 300
QY 301 YNYLSTDVSGCTVLCPLHNOEVTABDGTORCEKSGPCARVCYGLGMEHLREVRATVSAN 360
Db 301 YNYLSTDVSGCTVLCPLHNOEVTABDGTORCEKSGPCARVCYGLGMEHLREVRATVSAN 360
QY 361 IOEFAGCKKIFGSLAFBPESFDGDPASNTAPLOPBOLQVFETLEBITGYLIISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFBPESFDGDPASNTAPLOPBOLQVFETLEBITGYLIISAMPDSL 420
QY 421 DLSVFNQVIRGRILHNGAYSILTQIGISWLGIRSLREIGSGIALHNNHTLCFYHTV 480
Db 421 DLSVFNQVIRGRILHNGAYSILTQIGISWLGIRSLREIGSGIALHNNHTLCFYHTV 480
QY 481 PMDOLFRPHQALHTANRPEDECVBGLACHQLCARHGCHGPGPTQCVCNSQFLRQCEC 540
Db 481 PMDOLFRPHQALHTANRPEDECVBGLACHQLCARHGCHGPGPTQCVCNSQFLRQCEC 540
QY 541 VEEBCVIOGLPREYVNAHCLPCHBECOPQNGSVTCFSEPADQCVACAHYKDPFCVARC 600
Db 541 VEEBCVIOGLPREYVNAHCLPCHBECOPQNGSVTCFSEPADQCVACAHYKDPFCVARC 600
QY 601 PSQVPRDLSYMPIKFPBEBGACQPCPINCHSCVDLDDKCPAORASPLTISIAYVG 660
Db 601 PSQVPRDLSYMPIKFPBEBGACQPCPINCHSCVDLDDKCPAORASPLTISIAYVG 660
QY 661 ILVVVVLGVNFGIILKRRQOKIRKTYMRLLQETLVEPLTPSGAMPQAQMRILKETEL 720
Db 661 ILVVVVLGVNFGIILKRRQOKIRKTYMRLLQETLVEPLTPSGAMPQAQMRILKETEL 720
QY 721 RKVVLSSGAFGTVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELIDEAYVNAVGSP 780
Db 721 RKVVLSSGAFGTVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELIDEAYVNAVGSP 780
QY 781 YVSRLLGICLTSTQVLTQMLPVGCLLDHVRNBRRLSODLLMCMQIAGMSYLEVR 840
Db 781 YVSRLLGICLTSTQVLTQMLPVGCLLDHVRNBRRLSODLLMCMQIAGMSYLEVR 840
QY 841 LVHRDLARNVLYKSPNHVKTDFGLARLLDIDEFEYADGGKYVIKMMALESIIRRF 900
Db 841 LVHRDLARNVLYKSPNHVKTDFGLARLLDIDEFEYADGGKYVIKMMALESIIRRF 900
QY 901 HQSDVWSYGVTVWELMTFGAKPYDGI PAKEIPDLLEKGBRLPQPICTIDVYMIWVKWM 960
Db 901 HQSDVWSYGVTVWELMTFGAKPYDGI PAKEIPDLLEKGBRLPQPICTIDVYMIWVKWM 960
QY 961 IDSERRPRPRVLSFSSMAARDPQRFVYQNEIDLCPASPLOSTFFRSILLEDDMDGLDVA 1020
Db 961 IDSERRPRPRVLSFSSMAARDPQRFVYQNEIDLCPASPLOSTFFRSILLEDDMDGLDVA 1020
QY 1021 EBYLVPQOGFFCPDPAPGAGVWHHRHSSSTRSGGDLTLGLBPSSEBAPRSLAPSEG 1080
Db 1021 EBYLVPQOGFFCPDPAPGAGVWHHRHSSSTRSGGDLTLGLBPSSEBAPRSLAPSEG 1080
QY 1081 AGSDVFDGLMGAAKGLQSLPTHDPSPLQRYSEDPVFLPSETDGYVAPLTCSQPPEV 1140

```

```

Db 1081 AGSDVFDGLMGAAKGLQSLPTHDPSPLQRYSEDPVFLPSETDGYVAPLTCSQPPEV 1140
QY 1141 NQDVRPOPSPRRGRLPAARPAQATLERPTLLSPGKNGVQVDFAFGAVENPEYLRPQ 1200
Db 1141 NQDVRPOPSPRRGRLPAARPAQATLERPTLLSPGKNGVQVDFAFGAVENPEYLRPQ 1200
QY 1201 GGAAPQHPHPPAFSFAFDNLVYWDQDPPBRGAPSTFKGTPTAENPEYLGLDVVP 1255
Db 1201 GGAAPQHPHPPAFSFAFDNLVYWDQDPPBRGAPSTFKGTPTAENPEYLGLDVVP 1255

RESULT 8
AAU77114
ID AAU77114 standard; protein; 1255 AA.
XX
AC AAU77114;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human Her-2/neu polypeptide.
XX
KW Human; Her-2/neu; cytostatic; haematological malignancy; CML;
KW acute myelogenous leukaemia; AML; chronic myelogenous leukaemia; CLL;
KW chronic lymphocytic leukaemia; myeloma; non-Hodgkin's lymphoma; MDS;
KW Hodgkin's lymphoma; T cell therapy.
XX
OS Homo sapiens.
XX
PN WO200213847-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US025408.
XX
PR 14-AUG-2000; 2000US-00638280.
PR 28-SEP-2000; 2000US-00675904.
XX
PA (CORI-) CORIXA CORP.
XX
PI Gaiger A, Cheever MA, Hand-Zimmermann S;
XX
DR WPI; 2002-280741/32.
XX
DR N-PSDB; ABK10730.
XX
PT Inhibiting hematological malignancy development by administering
PT polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide
PT encoding the polypeptide, or antigen presenting cells expressing the
PT polypeptide.
XX
PS Disclosure; Page 71-74; 74pp; English.
XX
CC The invention relates to a method for inhibiting development of
CC haematological malignancy in a patient by administering a polypeptide
CC comprising an immunogenic portion of Her-2/neu or a polynucleotide
CC encoding the polypeptide. Antigen presenting cells that express the
CC protein can also be administered. The sequences are used for inhibiting
CC development of haematological malignancy such as acute myelogenous
CC leukaemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic
CC leukaemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's
CC lymphoma. This sequence represents the human Her-2/neu polypeptide
XX
SQ Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 5; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELAALCRWGILLALLLPGAASVQVCTGTDMKRLPASPETHLDMRLRLYOGCQVVGNTL 60
Db 1 MELAALCRWGILLALLLPGAASVQVCTGTDMKRLPASPETHLDMRLRLYOGCQVVGNTL 60
QY 61 ELTYLPTNASISFLDIOEVQGYVLIANHVRQVPLQRLRIVRGQLFEDNYALAVLDNG 120

```

```

Db      |||
61      ELTYLPTNASLFLDIOIEGVGVLLAHNQVQVPLQRLIRVGTQLFEDNYALALVDNG 120
Qy      |||
121     DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLTORNPOLCYOPTIIMKDI FHKNNOLA 180
Db      |||
121     DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLTORNPOLCYOPTIIMKDI FHKNNOLA 180
Qy      |||
181     LTLITNRSRACHPCSPMKGSRGWSSESDCOSLTRTYCAGGACARCKGPLPTDCHEQC 240
Db      |||
181     LTLITNRSRACHPCSPMKGSRGWSSESDCOSLTRTYCAGGACARCKGPLPTDCHEQC 240
Qy      |||
241     AAGCTGPKRSDCLACLFHNSGICBLHCPALVYNTDTEESMPNBEGRYTFGASCVTACP 300
Db      |||
241     AAGCTGPKRSDCLACLFHNSGICBLHCPALVYNTDTEESMPNBEGRYTFGASCVTACP 300
Qy      |||
301     XNYLSTDVGSCTLVCPLNQVTAEDGTORCEKSPCARVCYGLGMEHLREVRVTSAN 360
Db      |||
301     XNYLSTDVGSCTLVCPLNQVTAEDGTORCEKSPCARVCYGLGMEHLREVRVTSAN 360
Qy      |||
361     IOEPAGCKKIFGSLAFLPESFDGDDPASNTAPLOPELOVFEETLEETGLYISAMPDSL 420
Db      |||
361     IOEPAGCKKIFGSLAFLPESFDGDDPASNTAPLOPELOVFEETLEETGLYISAMPDSL 420
Qy      |||
421     DLSVFQNLQVIRGRILHNGAVSLTLQGLGISWLGLRSLRELSGGLAIHNTHLCEVHTV 480
Db      |||
421     DLSVFQNLQVIRGRILHNGAVSLTLQGLGISWLGLRSLRELSGGLAIHNTHLCEVHTV 480
Qy      |||
481     PMDOLFRNPHQALLHTANRPEDECVGEGACIACIOLCARHGCWPGPTOCNCQOFLRGQEC 540
Db      |||
481     PMDOLFRNPHQALLHTANRPEDECVGEGACIACIOLCARHGCWPGPTOCNCQOFLRGQEC 540
Qy      |||
541     VEECVLQGLPREVYNARHCLPCHPECPONGSVTCFGEBAQCYACAHYKDPPECVAR 600
Db      |||
541     VEECVLQGLPREVYNARHCLPCHPECPONGSVTCFGEBAQCYACAHYKDPPECVAR 600
Qy      |||
601     PSQVPRDLSYMBIMKRPDEBGAQCPPICTHSCVDLDDKGPABGRASPLTSIIISAVVG 660
Db      |||
601     PSQVPRDLSYMBIMKRPDEBGAQCPPICTHSCVDLDDKGPABGRASPLTSIIISAVVG 660
Qy      |||
661     ILVVVVLGVVFGILLKRRQOKIRKTYMRLLQETELVEPLPSGAMPNOQWRILKEETEL 720
Db      |||
661     ILVVVVLGVVFGILLKRRQOKIRKTYMRLLQETELVEPLPSGAMPNOQWRILKEETEL 720
Qy      |||
721     RKVKVLGSGAFGTVKGIWIPDGENVKIPVAIKVIRENTSPKANKIIDEAVYMAVGVS 780
Db      |||
721     RKVKVLGSGAFGTVKGIWIPDGENVKIPVAIKVIRENTSPKANKIIDEAVYMAVGVS 780
Qy      |||
781     YVSRLLGICLTSTVQVLTQMLPYGCLLDHVRNENRGRLSGODLLNMCQIAKGMSTLEDR 840
Db      |||
781     YVSRLLGICLTSTVQVLTQMLPYGCLLDHVRNENRGRLSGODLLNMCQIAKGMSTLEDR 840
Qy      |||
841     LVHRDLAARNVLKSPNHYKITDPGLARLLIDETBYHADGKVPKIMMALESIIIRRP 900
Db      |||
841     LVHRDLAARNVLKSPNHYKITDPGLARLLIDETBYHADGKVPKIMMALESIIIRRP 900
Qy      |||
901     HOSDWWSGVTVWELMTFGAKPYDGI PAEIPDLLEKGRLLPOPPCTIDVVMIMVKCM 960
Db      |||
901     HOSDWWSGVTVWELMTFGAKPYDGI PAEIPDLLEKGRLLPOPPCTIDVVMIMVKCM 960
Qy      |||
961     IDSECRPRRELVSFSSRMARDPQRFVVIQNEIDLSPASPLDSTFYRSLLEDDMDGLDVA 1020
Db      |||
961     IDSECRPRRELVSFSSRMARDPQRFVVIQNEIDLSPASPLDSTFYRSLLEDDMDGLDVA 1020
Qy      |||
1021     EBYLVPQOGFCPPDPAFGAGVHHRHRSSTRSGGGDTLCLPSEBEEAPSPPLAPSEG 1080
Db      |||
1021     EBYLVPQOGFCPPDPAFGAGVHHRHRSSTRSGGGDTLCLPSEBEEAPSPPLAPSEG 1080
Qy      |||
1081     AGSDVFDGDLGMAKGLQSLPTHDSPLOQVSEDPVPLPSETTGUYAPLTCSQOPEV 1140
Db      |||
1081     AGSDVFDGDLGMAKGLQSLPTHDSPLOQVSEDPVPLPSETTGUYAPLTCSQOPEV 1140
Qy      |||
1141     NOPDVRPOPSPREGPLPAARPAAGATLERPKTLSBGKGVVDVAFGGAVERNPEYLTPQ 1200

```

```

Db      |||
1141     NOPDVRPOPSPREGPLPAARPAAGATLERPKTLSBGKGVVDVAFGGAVERNPEYLTPQ 1200
Qy      |||
1201     GGAAPQPPPAFSPAFDNLVYWDODPPRGAPSPSTFGKPTPAENPEYLGDPV 1255
Db      |||
1201     GGAAPQPPPAFSPAFDNLVYWDODPPRGAPSPSTFGKPTPAENPEYLGDPV 1255
RESULT 9
AAMS1143
ID AAMS1143 strand: protein; 1255 AA.
XX
AC AAMS1143;
XX
DT 17-JUN-2002 (first entry)
XX
DE Human Her-2/neu oncogene-encoded p185 glycoprotein.
XX
KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
KW tyrosine kinase; receptor; c-erbB2; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..653
FT /note="extracellular domain"
FT Domain 676..1255
FT /note="intracellular domain"
FT Domain 990..1255
FT /note="phosphorylation domain"
PN WO200212341-A2.
PD 14-FEB-2002.
PR 03-AUG-2001; 2001WO-US024283.
PR 03-AUG-2000; 2000US-00632507.
PA (CORI-) CORIYA CORP.
PA (SMK) SMITHLINE BECHAM BIOLOGICALS.
PI Cheever MA, Gheysen D;
XX
DR WPI: 2002-241743/29.
DR N-PSDB; ABA92250.
XX
PT Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT or enhancing an immune response to the protein, has Her-2/neu
PT extracellular domain fused to Her-2/neu intracellular or phosphorylation
PT domain.
XX
PS Claim 68; Fig 7; 141pp; English.
XX
CC The present sequence is that of human Her-2/neu (p185 glycoprotein or c-
CC erbB2), an oncogenic self-protein and target for anti-cancer vaccines.
CC The Her-2/neu gene is amplified and p185 is overexpressed in a variety of
CC cancers, including breast, ovarian, colon, lung and prostate cancer. Her-
CC 2/neu is a member of the tyrosine kinase family of receptor-like
CC glycoproteins. It comprises an extracellular domain with homology to the
CC epidermal growth factor receptor (EGFR), a highly hydrophobic
CC transmembrane domain and a C-terminal intracellular domain that also
CC shows homology to EGFR. Its overexpression correlates with a poor
CC prognosis in breast and ovarian cancers. The invention provides Her-2/neu
CC fusion proteins, nucleic acids encoding them, viral vectors, and vaccines
CC comprising the fusion proteins or nucleic acid molecules. In preferred
CC fusion proteins, the extracellular domain of a Her-2/neu protein is fused
CC to a Her-2/neu intracellular domain or phosphorylation domain (or its
CC Delaap fragment). An immune response to Her-2/neu protein is elicited or
CC enhanced by administering the fusion protein in the form of a vaccine, or
CC by transfecting cells of an animal ex vivo with a nucleic acid encoding
CC the fusion protein, and delivering the transfect cells to the animal.
CC The fusion proteins, nucleic acids, and isolated specific T-cells are
CC useful for inhibiting the development of a cancer, especially breast,

```

ovarian, colon, lung or prostate cancer in a patient. T cells that specifically react with a Her-2/neu fusion protein can be used to remove tumour cells from a sample in order to inhibit the development of cancer in a patient

Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 5; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELAALCRWGLLALILPPGAASSTQVCTGTMKLRIPASPEITHLDMRLHYGCGQVQGNL 60
DB 1 MELAALCRWGLLALILPPGAASSTQVCTGTMKLRIPASPEITHLDMRLHYGCGQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVGVYLAHNQVQVPLQRLRYRGTQLFEDNIALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVGVYLAHNQVQVPLQRLRYRGTQLFEDNIALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLTEILKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLTEILKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
QY 121 DPLNNTTPTVGTASPGGLRELQRLTEILKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLTEILKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
QY 181 LTLIDTNSRACHPCSPWCKSGRCWSESSDCQSLTRTVCAAGCARCKGPIPTDCHEQC 240
DB 181 LTLIDTNSRACHPCSPWCKSGRCWSESSDCQSLTRTVCAAGCARCKGPIPTDCHEQC 240
QY 241 AAGCTPRGHSQCLALHFNHSGICELHCPALVTNTDFESPNRNGRYTGASCVTACP 300
DB 241 AAGCTPRGHSQCLALHFNHSGICELHCPALVTNTDFESPNRNGRYTGASCVTACP 300
QY 241 AAGCTPRGHSQCLALHFNHSGICELHCPALVTNTDFESPNRNGRYTGASCVTACP 300
DB 241 AAGCTPRGHSQCLALHFNHSGICELHCPALVTNTDFESPNRNGRYTGASCVTACP 300
QY 301 YNYLSTDVSGCTIVCPPLHNOEVTAEQTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
DB 301 YNYLSTDVSGCTIVCPPLHNOEVTAEQTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
QY 301 YNYLSTDVSGCTIVCPPLHNOEVTAEQTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
DB 301 YNYLSTDVSGCTIVCPPLHNOEVTAEQTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
QY 361 IQEPAGCKKIRGSLAFLPESPDGPASNTAPLQEPOLQVFTLEITIGLYLISAMPDLP 420
DB 361 IQEPAGCKKIRGSLAFLPESPDGPASNTAPLQEPOLQVFTLEITIGLYLISAMPDLP 420
QY 361 IQEPAGCKKIRGSLAFLPESPDGPASNTAPLQEPOLQVFTLEITIGLYLISAMPDLP 420
DB 361 IQEPAGCKKIRGSLAFLPESPDGPASNTAPLQEPOLQVFTLEITIGLYLISAMPDLP 420
QY 421 DLSVFQNLQVIRGRILHNHGAYSLTQGLGISWLGSRISRELSGSLALHNHNTHLCEVHTV 480
DB 421 DLSVFQNLQVIRGRILHNHGAYSLTQGLGISWLGSRISRELSGSLALHNHNTHLCEVHTV 480
QY 421 DLSVFQNLQVIRGRILHNHGAYSLTQGLGISWLGSRISRELSGSLALHNHNTHLCEVHTV 480
DB 421 DLSVFQNLQVIRGRILHNHGAYSLTQGLGISWLGSRISRELSGSLALHNHNTHLCEVHTV 480
QY 481 PMDQLFRNPQALHTANRPEDECEVGEGLACHQLCARGHCWGPPTQCVNSQFLRGQEC 540
DB 481 PMDQLFRNPQALHTANRPEDECEVGEGLACHQLCARGHCWGPPTQCVNSQFLRGQEC 540
QY 481 PMDQLFRNPQALHTANRPEDECEVGEGLACHQLCARGHCWGPPTQCVNSQFLRGQEC 540
DB 481 PMDQLFRNPQALHTANRPEDECEVGEGLACHQLCARGHCWGPPTQCVNSQFLRGQEC 540
QY 541 VEECEVLTQGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCYACHYKDPPECVARC 600
DB 541 VEECEVLTQGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCYACHYKDPPECVARC 600
QY 541 VEECEVLTQGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCYACHYKDPPECVARC 600
DB 541 VEECEVLTQGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCYACHYKDPPECVARC 600
QY 601 PSGVPEDLSTYPIWKEPDEBGACOPCINCHSCVDLDDKGPAPORASPLTSTISAVVG 660
DB 601 PSGVPEDLSTYPIWKEPDEBGACOPCINCHSCVDLDDKGPAPORASPLTSTISAVVG 660
QY 601 PSGVPEDLSTYPIWKEPDEBGACOPCINCHSCVDLDDKGPAPORASPLTSTISAVVG 660
DB 601 PSGVPEDLSTYPIWKEPDEBGACOPCINCHSCVDLDDKGPAPORASPLTSTISAVVG 660
QY 661 ILVVVLTGVVFGILIKRQOKIRKTYRRLQETELVEPLTPSGAMPNQAQMRILKETEL 720
DB 661 ILVVVLTGVVFGILIKRQOKIRKTYRRLQETELVEPLTPSGAMPNQAQMRILKETEL 720
QY 661 ILVVVLTGVVFGILIKRQOKIRKTYRRLQETELVEPLTPSGAMPNQAQMRILKETEL 720
DB 661 ILVVVLTGVVFGILIKRQOKIRKTYRRLQETELVEPLTPSGAMPNQAQMRILKETEL 720
QY 721 RKVVLVSGAGFTVYKGIWIPGGENVKIPVAIKVIRENTSPKANKEIIDEAIVMAGVGSP 780
DB 721 RKVVLVSGAGFTVYKGIWIPGGENVKIPVAIKVIRENTSPKANKEIIDEAIVMAGVGSP 780
QY 721 RKVVLVSGAGFTVYKGIWIPGGENVKIPVAIKVIRENTSPKANKEIIDEAIVMAGVGSP 780
DB 721 RKVVLVSGAGFTVYKGIWIPGGENVKIPVAIKVIRENTSPKANKEIIDEAIVMAGVGSP 780
QY 781 YVSRLLGICLSTVQLVQMLPMPYGLLDHVRENRRGLSODLLNMCQIAGKMSYLEVVR 840
DB 781 YVSRLLGICLSTVQLVQMLPMPYGLLDHVRENRRGLSODLLNMCQIAGKMSYLEVVR 840
QY 781 YVSRLLGICLSTVQLVQMLPMPYGLLDHVRENRRGLSODLLNMCQIAGKMSYLEVVR 840
DB 781 YVSRLLGICLSTVQLVQMLPMPYGLLDHVRENRRGLSODLLNMCQIAGKMSYLEVVR 840
QY 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETEHADGGKVPKIMMALESILRRFT 900
DB 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETEHADGGKVPKIMMALESILRRFT 900
QY 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETEHADGGKVPKIMMALESILRRFT 900
DB 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETEHADGGKVPKIMMALESILRRFT 900
QY 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLKEGRLPQPICTIDVYMLVKCM 960
DB 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLKEGRLPQPICTIDVYMLVKCM 960

```

```

DB 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLKEGRLPQPICTIDVYMLVKCM 960
QY 961 IDSECRPRELVESEFARMARDPQRFVYIQNEDLGPA SPLDSTFYRSLLEDMDGLVDA 1020
DB 961 IDSECRPRELVESEFARMARDPQRFVYIQNEDLGPA SPLDSTFYRSLLEDMDGLVDA 1020
QY 1021 EBYLVPOQGFPCDPAPAGAMVHHRRSSSTRSGGDLTYGLBSEBARSPPLABSEG 1080
DB 1021 EBYLVPOQGFPCDPAPAGAMVHHRRSSSTRSGGDLTYGLBSEBARSPPLABSEG 1080
QY 1021 EBYLVPOQGFPCDPAPAGAMVHHRRSSSTRSGGDLTYGLBSEBARSPPLABSEG 1080
DB 1021 EBYLVPOQGFPCDPAPAGAMVHHRRSSSTRSGGDLTYGLBSEBARSPPLABSEG 1080
QY 1081 AGSDVPFGDGLGMAKAGLSLPTHDSPLOKYSBDPTVPLSEYDGYVAPLTCSPQPEYV 1140
DB 1081 AGSDVPFGDGLGMAKAGLSLPTHDSPLOKYSBDPTVPLSEYDGYVAPLTCSPQPEYV 1140
QY 1141 NQPDVRRQPPSPRSGPLPAARPAATLERPKTSLSPKNGVYKDYAFAGAVENPEYLTPO 1200
DB 1141 NQPDVRRQPPSPRSGPLPAARPAATLERPKTSLSPKNGVYKDYAFAGAVENPEYLTPO 1200
QY 1201 GGAAPQHPHPAPSPAPDNLTYWDQDPPERCAPSTFKGPTTANPEYLGIDVPY 1255
DB 1201 GGAAPQHPHPAPSPAPDNLTYWDQDPPERCAPSTFKGPTTANPEYLGIDVPY 1255

```

```

RESULT 10
AAE24067
ID AAE24067 standard; protein, 1255 AA.
XX
AC AAE24067;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human Her-2 protein.
XX
KW Human; Her-2; epidermal growth factor receptor 2; infection; cancer;
hyperproliferative disorder; prophylaxis; inflammation; antisense;
tumour; gene therapy; phosphothiotate backbone.
XX
OS Homo sapiens.
XX
PN MO200222636-A1.
XX
PD 21-MAR-2002.
XX
PE 12-SEP-2001; 2001WO-US028572.
XX
PR 15-SEP-2000; 2000US-00663834.
XX
PA (ISIS-) ISIS PHARM INC.
PI Bennett CF, Cowseart LM;
PI WPI; 2002-471192/50.
PI N-PSDB; AAD38904.
PT Epidermal Growth Factor receptor, Her2, is useful for treating tumors
inflammation or to prevent infection in humans.
PS Example 13; Page 95-107; 116pp; English.
XX
CC The invention relates to antisense compounds targeted to a nucleic acid
molecule encoding Her2 (human Epidermal Growth Factor receptor 2) that
specifically hybridizes with and inhibits the expression of Her2.
CC Antisense compounds of the invention are used for treating diseases or
conditions associated with Her2 such as hyperproliferative disorders e.g.
lung, breast, gastric, oesophageal, colon, bladder, salivary, neural or
cardiac cancer. They are also useful prophylactically e.g. to prevent or
delay infection, inflammation and tumour formation. The invention is also
used in gene therapy. The present sequence is human Her-2 protein
XX
SQ Sequence 1255 AA;

```

Query Match 100.0%; Score 6815; DB 5; Length 1255;



Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 METALACRGLLALLLPRAASTVCTGDMKLRPLASPEHLDMRLHYOCQVVGNTL 60
DB 1 METALACRGLLALLLPRAASTVCTGDMKLRPLASPEHLDMRLHYOCQVVGNTL 60
QY 61 ELTYLPTNASLFLQDIOEVGVVLIANQVQVPLQRLIRVGTQLFEDNALAVLNG 120
DB 61 ELTYLPTNASLFLQDIOEVGVVLIANQVQVPLQRLIRVGTQLFEDNALAVLNG 120
QY 121 DELNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTILMKDIFKNNQLA 180
DB 121 DELNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTILMKDIFKNNQLA 180
QY 181 LTLIDTNSRACHPCSPMKSGRCWSESEDCQSLTRTYCAGGCARCKGPLPTDCHEQC 240
DB 181 LTLIDTNSRACHPCSPMKSGRCWSESEDCQSLTRTYCAGGCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNPEGRTYFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNPEGRTYFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREAVTSAN 360
DB 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREAVTSAN 360
QY 361 IOEPFGCKKIFGSLAFPLPESFPGDPSANAPLOPQLOVFELEETLGYLISAMPDILP 420
DB 361 IOEPFGCKKIFGSLAFPLPESFPGDPSANAPLOPQLOVFELEETLGYLISAMPDILP 420
QY 421 DLVSQNLQVIRGRILHNGAVSLTQGLGISWGLRSLRELSGSLIHNTHLFCVTV 480
DB 421 DLVSQNLQVIRGRILHNGAVSLTQGLGISWGLRSLRELSGSLIHNTHLFCVTV 480
QY 481 FWDQLFRNPQALHTANRPEDECVBGLACHQLCARHCWPGPTQCVCNCSQPLRGQEC 540
DB 481 FWDQLFRNPQALHTANRPEDECVBGLACHQLCARHCWPGPTQCVCNCSQPLRGQEC 540
QY 541 VEBECVLOGLPREYVNAHCLPCHPECOPQNSVTCFGEADQCYACAHYKPPFCVARC 600
DB 541 VEBECVLOGLPREYVNAHCLPCHPECOPQNSVTCFGEADQCYACAHYKPPFCVARC 600
QY 601 PSGVAPDLSYMPIMKFPDEBEGACOPCINCHSCVDLDDKGPAPORASPLTSIIISAVYG 660
DB 601 PSGVAPDLSYMPIMKFPDEBEGACOPCINCHSCVDLDDKGPAPORASPLTSIIISAVYG 660
QY 661 ILLVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQAKRIKETEL 720
DB 661 ILLVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQAKRIKETEL 720
QY 721 RKVKVLGSGAGFTVYKGIWIPGSENVKIPVAIKVIRENTSPKANIELDEAVYMAVGSP 780
DB 721 RKVKVLGSGAGFTVYKGIWIPGSENVKIPVAIKVIRENTSPKANIELDEAVYMAVGSP 780
QY 781 YVSRLLGI CLTSTVOLVQLMPYGLLDHVRNRRGLSODLLNMCQIAKMSLTEDVR 840
DB 781 YVSRLLGI CLTSTVOLVQLMPYGLLDHVRNRRGLSODLLNMCQIAKMSLTEDVR 840
QY 841 LVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETEHYADGGKVDIKWMALESILRRRT 900
DB 841 LVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETEHYADGGKVDIKWMALESILRRRT 900
QY 901 HQSDWVSIGVYTWELMTFGAKPYDGI PAKEIDLLLEKGRLLPQPICTIDVYMIWVKCM 960
DB 901 HQSDWVSIGVYTWELMTFGAKPYDGI PAKEIDLLLEKGRLLPQPICTIDVYMIWVKCM 960
QY 961 IDSECRPPRRLVSEFSRMAKDPORFVYQNEDELPAASPJLSTFYSLLLEDDMDGLVDA 1020
DB 961 IDSECRPPRRLVSEFSRMAKDPORFVYQNEDELPAASPJLSTFYSLLLEDDMDGLVDA 1020
QY 1021 BEYLVPQOQFPCPDPAFGAGVNHHRHSSTRSGGDLTLGLBSEBEPAPRSLAPSEG 1080
DB 1021 BEYLVPQOQFPCPDPAFGAGVNHHRHSSTRSGGDLTLGLBSEBEPAPRSLAPSEG 1080

```

```

DB 1021 BEYLVPQOQFPCPDPAFGAGVNHHRHSSTRSGGDLTLGLBSEBEPAPRSLAPSEG 1080
QY 1081 AGSDVFDGLGWAAGKGLQSLPTHPSPLOQRYSEDPVPLPSETDGYVAPLCSPOPEVY 1140
DB 1081 AGSDVFDGLGWAAGKGLQSLPTHPSPLOQRYSEDPVPLPSETDGYVAPLCSPOPEVY 1140
QY 1141 NOPDVRPQPPSPREGPLPAARPAATLERPKTLSPGKGVVVDVAFGAVENPEYLTPQ 1200
DB 1141 NOPDVRPQPPSPREGPLPAARPAATLERPKTLSPGKGVVVDVAFGAVENPEYLTPQ 1200
QY 1201 GGAAPQHPPPAFSPAFDNLTYWDDPPBRGAPBSPSTFKGTPPAENPEYLGLDVPV 1255
DB 1201 GGAAPQHPPPAFSPAFDNLTYWDDPPBRGAPBSPSTFKGTPPAENPEYLGLDVPV 1255

RESULT 11
ABR43687
ID ABR43687 standard; protein; 1255 AA.
XX
AC ABR43687;
XX
DT 29-JUL-2003 (first entry)
XX
DE Human c-erb-B-2 (ErbB-2) protein SEQ ID NO:4.
XX
KW Human; hepatocellular carcinoma associated protein; HCAP; ErbB-2;
KW c-erb-B-2; cytostatic; nephrotropic; ophthalmological; antiproliferative;
KW antirheumatic; antiarthritic; antiinflammatory; HCAP-associated disease;
KW ErbB-2 associated disease; cell hyperproliferative disease; cancer;
KW cell proliferative disease; benign proliferative disease; psoriasis;
KW glomerulonephritis; neurofibromatosis; glaucoma; rheumatoid arthritis;
KW inflammatory bowel disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..21
FT /label= signal
FT 22..1255
FT /label= C-erb-B-2
FT /note= "also known as ErbB-2"
XX
PN WO2003025228-A1.
XX
PD 27-MAR-2003.
XX
PF 18-SEP-2002; 2002WO-US029577.
XX
PR 18-SEP-2001; 2001US-0323210P.
XX
PR 09-NOV-2001; 2001US-0323250P.
XX
PA (PROT-) PROTEOLOGICS INC.
XX
PI Reiss Y, Alroy I;
XX
DR MPI; 2003-363157/34.
XX
DR N-PSDB; ACC69999.
XX
PT Treating a hepatocellular carcinoma associated protein (HCAP)-associated
PT disease, e.g. cancer or benign proliferative disease, comprises
PT administering an agent that inhibits the interaction between HCAP and
PT HCAP binding partner.
XX
PS Claim 37; Fig 5C; 59pp; English.
XX
CC The present invention describes a method for treating a hepatocellular
CC carcinoma associated protein (HCAP)-associated disease in a subject which
CC comprises administering to the subject an agent that inhibits the
CC interaction between HCAP and an HCAP binding partner. Also described is a
CC method for treating an ErbB-2 associated disease in a subject which
CC comprises administering to the subject an agent that decreases the level
CC and/or activity of HCAP in cells of a subject. HCAP has cytostatic,
CC nephrotropic, ophthalmological, antiproliferative, antirheumatic,

```

CC antiarthritic and antiinflammatory activities. The methods can be used  
 CC for treating a HCAP-associated disease or an ErbB-2 associated disease,  
 CC e.g. a cell (hyper)proliferative disease such as cancer, or a benign  
 CC proliferative disease such as glomerulonephritis, neurofibromatosis,  
 CC glaucoma, psoriasis, rheumatoid arthritis or inflammatory bowel disease.  
 CC The methods can be used for treating a proliferative disorder, e.g.  
 CC breast cancer, ErbB-2-related breast cancer or taxol-resistant breast  
 CC cancer. The present sequence represents human c-erb-B-2 (also known as  
 CC ErbB-2) from the present invention

CC  
 XX Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 6; Length 1255;

Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCWGILLALLPPGAASVQVCTGTDMLRLPASPEHLDMLRHLVGGCQVQGNL 60  
 DB 1 MELAALCWGILLALLPPGAASVQVCTGTDMLRLPASPEHLDMLRHLVGGCQVQGNL 60  
 QY 61 ELTYLPTNASISFLQDIOEVQVYLIAHNOVRQVPLQSLRIYRGTLFEDNYALAVLDNG 120  
 DB 61 ELTYLPTNASISFLQDIOEVQVYLIAHNOVRQVPLQSLRIYRGTLFEDNYALAVLDNG 120  
 QY 121 DPLNNTTPTVGTASPGGLRELQLRSILTEILKGGVLIQRNPOLCYOPTILMKDIFHKNNQLA 180  
 DB 121 DPLNNTTPTVGTASPGGLRELQLRSILTEILKGGVLIQRNPOLCYOPTILMKDIFHKNNQLA 180  
 QY 121 DPLNNTTPTVGTASPGGLRELQLRSILTEILKGGVLIQRNPOLCYOPTILMKDIFHKNNQLA 180  
 DB 121 DPLNNTTPTVGTASPGGLRELQLRSILTEILKGGVLIQRNPOLCYOPTILMKDIFHKNNQLA 180  
 QY 181 LTLIDTNSRACHPSCPMCKSGRCSGSESDCOSLTRVVCAGGACRCGRLPTDCCHEOC 240  
 DB 181 LTLIDTNSRACHPSCPMCKSGRCSGSESDCOSLTRVVCAGGACRCGRLPTDCCHEOC 240  
 QY 181 LTLIDTNSRACHPSCPMCKSGRCSGSESDCOSLTRVVCAGGACRCGRLPTDCCHEOC 240  
 DB 181 LTLIDTNSRACHPSCPMCKSGRCSGSESDCOSLTRVVCAGGACRCGRLPTDCCHEOC 240  
 QY 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVTYTFPESMPNDEGRYFGASCVTACP 300  
 DB 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVTYTFPESMPNDEGRYFGASCVTACP 300  
 QY 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVTYTFPESMPNDEGRYFGASCVTACP 300  
 DB 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVTYTFPESMPNDEGRYFGASCVTACP 300  
 QY 301 YNYLSTDVGSCTLYVCPILHNOEVTADGTQRCCKSKPCARVCYGLGMEHLREVAVTISAN 360  
 DB 301 YNYLSTDVGSCTLYVCPILHNOEVTADGTQRCCKSKPCARVCYGLGMEHLREVAVTISAN 360  
 QY 301 YNYLSTDVGSCTLYVCPILHNOEVTADGTQRCCKSKPCARVCYGLGMEHLREVAVTISAN 360  
 DB 301 YNYLSTDVGSCTLYVCPILHNOEVTADGTQRCCKSKPCARVCYGLGMEHLREVAVTISAN 360  
 QY 361 IOEFAGCKKIFGSLAFLESFPGDPAANTAPLQEPOLQVFETLEETITGYLISAMPDLP 420  
 DB 361 IOEFAGCKKIFGSLAFLESFPGDPAANTAPLQEPOLQVFETLEETITGYLISAMPDLP 420  
 QY 361 IOEFAGCKKIFGSLAFLESFPGDPAANTAPLQEPOLQVFETLEETITGYLISAMPDLP 420  
 DB 361 IOEFAGCKKIFGSLAFLESFPGDPAANTAPLQEPOLQVFETLEETITGYLISAMPDLP 420  
 QY 421 DLSVQNTQVIRGRILIHNGAYSLTLQGLISWLGRLSRELGSGLALIHNTHLCFVHTV 480  
 DB 421 DLSVQNTQVIRGRILIHNGAYSLTLQGLISWLGRLSRELGSGLALIHNTHLCFVHTV 480  
 QY 421 DLSVQNTQVIRGRILIHNGAYSLTLQGLISWLGRLSRELGSGLALIHNTHLCFVHTV 480  
 DB 421 DLSVQNTQVIRGRILIHNGAYSLTLQGLISWLGRLSRELGSGLALIHNTHLCFVHTV 480  
 QY 481 FMDQLFRNPQALHTANRPEDECVSEGLACHQLCARGHCWGPPTQCVCNCSQFLRGQC 540  
 DB 481 FMDQLFRNPQALHTANRPEDECVSEGLACHQLCARGHCWGPPTQCVCNCSQFLRGQC 540  
 QY 481 FMDQLFRNPQALHTANRPEDECVSEGLACHQLCARGHCWGPPTQCVCNCSQFLRGQC 540  
 DB 481 FMDQLFRNPQALHTANRPEDECVSEGLACHQLCARGHCWGPPTQCVCNCSQFLRGQC 540  
 QY 541 VEECVVLOGLPREYVNAHCLPCHPECQPNQSVTCFEPBADQCAAHYDPPFCVARC 600  
 DB 541 VEECVVLOGLPREYVNAHCLPCHPECQPNQSVTCFEPBADQCAAHYDPPFCVARC 600  
 QY 541 VEECVVLOGLPREYVNAHCLPCHPECQPNQSVTCFEPBADQCAAHYDPPFCVARC 600  
 DB 541 VEECVVLOGLPREYVNAHCLPCHPECQPNQSVTCFEPBADQCAAHYDPPFCVARC 600  
 QY 601 PSGVVPDLSYMPYKPEPDEGACQPCPINCSTSCVDLDDKGPAPORASPLTISAVYG 660  
 DB 601 PSGVVPDLSYMPYKPEPDEGACQPCPINCSTSCVDLDDKGPAPORASPLTISAVYG 660  
 QY 601 PSGVVPDLSYMPYKPEPDEGACQPCPINCSTSCVDLDDKGPAPORASPLTISAVYG 660  
 DB 601 PSGVVPDLSYMPYKPEPDEGACQPCPINCSTSCVDLDDKGPAPORASPLTISAVYG 660  
 QY 661 ILVVVLGVGFILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQAQMRILKETEL 720  
 DB 661 ILVVVLGVGFILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQAQMRILKETEL 720  
 QY 661 ILVVVLGVGFILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQAQMRILKETEL 720  
 DB 661 ILVVVLGVGFILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQAQMRILKETEL 720  
 QY 721 RKVVLGSGAGFTYKGIWIPGGEVVKIPVAIKVRENTSPANKIIDEAYVAGVSP 780  
 DB 721 RKVVLGSGAGFTYKGIWIPGGEVVKIPVAIKVRENTSPANKIIDEAYVAGVSP 780  
 QY 721 RKVVLGSGAGFTYKGIWIPGGEVVKIPVAIKVRENTSPANKIIDEAYVAGVSP 780  
 DB 721 RKVVLGSGAGFTYKGIWIPGGEVVKIPVAIKVRENTSPANKIIDEAYVAGVSP 780  
 QY 781 YVSRLLGICLTSTVQLVQMLPYGCLDHPVRENRRGLSODLLNMCQIJAQMSYLEYVR 840  
 DB 781 YVSRLLGICLTSTVQLVQMLPYGCLDHPVRENRRGLSODLLNMCQIJAQMSYLEYVR 840  
 QY 781 YVSRLLGICLTSTVQLVQMLPYGCLDHPVRENRRGLSODLLNMCQIJAQMSYLEYVR 840  
 DB 781 YVSRLLGICLTSTVQLVQMLPYGCLDHPVRENRRGLSODLLNMCQIJAQMSYLEYVR 840  
 QY 841 LVHRDLAARNVLYKSPNHVKITDQGLARLLIDETEHADGGKVPDKWMALESILRRFT 900  
 DB 841 LVHRDLAARNVLYKSPNHVKITDQGLARLLIDETEHADGGKVPDKWMALESILRRFT 900

DB 841 LVHRDLAARNVLYKSPNHVKITDQGLARLLIDETEHADGGKVPDKWMALESILRRFT 900  
 QY 901 HQSVMGVTWMLMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYMIWVKCM 960  
 DB 901 HQSVMGVTWMLMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYMIWVKCM 960  
 QY 901 HQSVMGVTWMLMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYMIWVKCM 960  
 DB 901 HQSVMGVTWMLMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYMIWVKCM 960  
 QY 961 IDSECRPRFRELVEFSRMAADPQRFVVIQNEIDGAPASPLDSTFYRSLLEDDMDGLVDA 1020  
 DB 961 IDSECRPRFRELVEFSRMAADPQRFVVIQNEIDGAPASPLDSTFYRSLLEDDMDGLVDA 1020  
 QY 961 IDSECRPRFRELVEFSRMAADPQRFVVIQNEIDGAPASPLDSTFYRSLLEDDMDGLVDA 1020  
 DB 961 IDSECRPRFRELVEFSRMAADPQRFVVIQNEIDGAPASPLDSTFYRSLLEDDMDGLVDA 1020  
 QY 1021 EEYLVPOQGFPCPPAAGAGMTHRRSSSTRSGGDLTLGLPSEEARSPLABSEG 1080  
 DB 1021 EEYLVPOQGFPCPPAAGAGMTHRRSSSTRSGGDLTLGLPSEEARSPLABSEG 1080  
 QY 1021 EEYLVPOQGFPCPPAAGAGMTHRRSSSTRSGGDLTLGLPSEEARSPLABSEG 1080  
 DB 1021 EEYLVPOQGFPCPPAAGAGMTHRRSSSTRSGGDLTLGLPSEEARSPLABSEG 1080  
 QY 1081 AGSDVFDGDLGMGAAGKQSLPTHDPSPLQRYSDPVLPLSFEDGYVAPLTCSPOPEYV 1140  
 DB 1081 AGSDVFDGDLGMGAAGKQSLPTHDPSPLQRYSDPVLPLSFEDGYVAPLTCSPOPEYV 1140  
 QY 1141 NQPDVRRQPSPRBGPPLAARPAATLERPKTSLPGKNGVVKDVFARFAGAVENPEYLTPQ 1200  
 DB 1141 NQPDVRRQPSPRBGPPLAARPAATLERPKTSLPGKNGVVKDVFARFAGAVENPEYLTPQ 1200  
 QY 1141 NQPDVRRQPSPRBGPPLAARPAATLERPKTSLPGKNGVVKDVFARFAGAVENPEYLTPQ 1200  
 DB 1141 NQPDVRRQPSPRBGPPLAARPAATLERPKTSLPGKNGVVKDVFARFAGAVENPEYLTPQ 1200  
 QY 1201 GGAAPQHPHPPAPESPAEDNLYYWDQDPPERKGAPESTKGTPTAANPEYLGLDVPV 1255  
 DB 1201 GGAAPQHPHPPAPESPAEDNLYYWDQDPPERKGAPESTKGTPTAANPEYLGLDVPV 1255  
 QY 1201 GGAAPQHPHPPAPESPAEDNLYYWDQDPPERKGAPESTKGTPTAANPEYLGLDVPV 1255  
 DB 1201 GGAAPQHPHPPAPESPAEDNLYYWDQDPPERKGAPESTKGTPTAANPEYLGLDVPV 1255

RESULT 12  
 ABR82066  
 ID ABR82066 standard; protein: 1255 AA.  
 AC ABR82066;  
 XX 23-SEP-2003 (first entry)  
 DT 23-SEP-2003 (first entry)  
 XX Human Her2/neu amino acid sequence SEQ ID NO:1.  
 DE Human Her2/neu; Her2/neu target antigen; immune response; cytostatic;  
 XX Human; Her2/neu; Her2/neu target antigen; immune response; cytostatic;  
 XX immunostimulant; vaccine; cancer.  
 OS Homo sapiens.  
 XX MO2003055439-A2.  
 PN 10-JUL-2003.  
 PD 10-JUL-2003.  
 XX 18-JUL-2002; 2002MO-US022975.  
 PF 18-JUL-2002; 2002MO-US022975.  
 XX 18-JUL-2001; 2001US-0306250P.  
 PR 18-JUL-2001; 2001US-0306250P.  
 XX (REGC) UNIV CALIFORNIA.  
 PA Nelson EL;  
 PI WPI; 2003-569400/53.  
 DR WPI; 2003-569400/53.  
 XX New Her2/neu target antigens and polynucleotides encoding them, useful  
 XX for stimulating immunoeffector cells or an immune response in a subject  
 PT against cancer cells expressing Her2/neu, and for vaccination strategies.  
 XX Claim 29; Fig 1A-B; 83p; English.  
 PS The present invention describes a recombinant polynucleotide (1)  
 CC comprising a first nucleotide sequence encoding a Her2/neu target antigen  
 CC consisting of an amino acid sequence corresponding to: (a) amino acid  
 CC residues 634-683 or 606-683 operatively linked to amino acid residues  
 CC 1035-1255 of human Her2/neu (SEQ ID NO:1, ABR82066); or (b) amino acid  
 CC residues 635-685 or 608-685 operatively linked to amino acid residues  
 CC 1037-1257 of rat Her2/neu (SEQ ID NO:2, ABR82067). (1) has cytostatic and  
 CC immunostimulant activities, and can be used in vaccines. The  
 CC polynucleotide encoding a Her2/neu target antigen can be used for  
 CC stimulating immunoeffector cells or an immune response in a subject

CC against cancer cells expressing Her2/neu. The encoded Her2/neu target  
 CC antigen polypeptide is useful as a tag to detect or isolate the fusion  
 CC protein containing the Her2/neu target an immune response stimulated by  
 CC the Her2/neu target antigen. The Her2/neu target antigen polynucleotides  
 CC and polypeptides are useful in genetic or polypeptide vaccination  
 CC strategies. The present sequence represents the human Her2/neu amino acid  
 CC sequence, which is used in the exemplification of the present invention  
 XX

XX Sequence 1255 AA;

Query Match 100.0%; Score 6015; DB 7; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELALCRKGLLALLPAGASTOVCTGDMKRLRASPETHLDMRLHYOCQVVGML 60  
 DB 1 MELALCRKGLLALLPAGASTOVCTGDMKRLRASPETHLDMRLHYOCQVVGML 60  
 QY 61 ELTYLPTNASLSFLQDIOEVQGVLLAHNQVRQVPLQRLIRVGTQLFEDNYALAVLDNG 120  
 DB 61 ELTYLPTNASLSFLQDIOEVQGVLLAHNQVRQVPLQRLIRVGTQLFEDNYALAVLDNG 120  
 QY 121 DELNATTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPQLCYODTILMKDIIFKXNQIA 180  
 DB 121 DELNATTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPQLCYODTILMKDIIFKXNQIA 180  
 QY 181 LTLITNRRACHPCSPMKSGRCWGESSEDCQSLTRTYCAGGACRCKPLPTDCHEC 240  
 DB 181 LTLITNRRACHPCSPMKSGRCWGESSEDCQSLTRTYCAGGACRCKPLPTDCHEC 240  
 QY 241 AAGCTPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNEGRTFGASCVTACP 300  
 DB 241 AAGCTPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNEGRTFGASCVTACP 300  
 QY 301 YNYLSTDVGSCTLVCPRLHNOEVTAEEDGTORCEKSKPCARVCYGLAMEHLREVAVTSAN 360  
 DB 301 YNYLSTDVGSCTLVCPRLHNOEVTAEEDGTORCEKSKPCARVCYGLAMEHLREVAVTSAN 360  
 QY 361 IOEPFAGCKKIFGSLAFLPESPFGDPAASNTAPLOPELOVFELEETLGTGLYSAMPDSL 420  
 DB 361 IOEPFAGCKKIFGSLAFLPESPFGDPAASNTAPLOPELOVFELEETLGTGLYSAMPDSL 420  
 QY 421 DLSVFQNLQVIRGRILHNGAVSLTLQGLISWLGRLSLRELSGGLAIHHNTHLCFVHTV 480  
 DB 421 DLSVFQNLQVIRGRILHNGAVSLTLQGLISWLGRLSLRELSGGLAIHHNTHLCFVHTV 480  
 QY 481 PMDQLFRNHQALLHTANRPEDECVGEGLAGHQLCARHGWPGPTQCVNCSQPLRGQEC 540  
 DB 481 PMDQLFRNHQALLHTANRPEDECVGEGLAGHQLCARHGWPGPTQCVNCSQPLRGQEC 540  
 QY 541 VEECVLQGLPREYVNAHCLPCHPECPQNSVTCFGEADQCVACAHYKDPFCVAC 600  
 DB 541 VEECVLQGLPREYVNAHCLPCHPECPQNSVTCFGEADQCVACAHYKDPFCVAC 600  
 QY 601 PSGVPRDISYMPIMKFPDEGACQPCPINCTHSCVDLDKGPABORASPLTISIAYVG 660  
 DB 601 PSGVPRDISYMPIMKFPDEGACQPCPINCTHSCVDLDKGPABORASPLTISIAYVG 660  
 QY 661 ILVVVLGVVFGILIKRQOKIRKTYMRLLQETELVEELPSGAMPQAOQRILKETEL 720  
 DB 661 ILVVVLGVVFGILIKRQOKIRKTYMRLLQETELVEELPSGAMPQAOQRILKETEL 720  
 QY 721 RRVKVLGSGAFGTVYKGIIPDGENVKIPVALKRENTSPANKIIDEAVYMGVGP 780  
 DB 721 RRVKVLGSGAFGTVYKGIIPDGENVKIPVALKRENTSPANKIIDEAVYMGVGP 780  
 QY 781 YSRLLGICLTSTVOLVTOLMPYGLLDHVRNRRGLSGODLLNMCQJAKGSLYEDVR 840  
 DB 781 YSRLLGICLTSTVOLVTOLMPYGLLDHVRNRRGLSGODLLNMCQJAKGSLYEDVR 840  
 QY 841 LVHRDLAARNVAVKSPNHVKITDPGLARLLIDETRYHADGKVPKIMALESIIRRP 900  
 DB 841 LVHRDLAARNVAVKSPNHVKITDPGLARLLIDETRYHADGKVPKIMALESIIRRP 900

QY 901 HOSDWSYGVYWEHMTFGAKPYDGI PARHIFDLLEKGERLPQPICTIDVYIMVCKM 960  
 DB 901 HOSDWSYGVYWEHMTFGAKPYDGI PARHIFDLLEKGERLPQPICTIDVYIMVCKM 960  
 QY 961 IDSECRPRRELVSFSSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSLLEDMDGLVDA 1020  
 DB 961 IDSECRPRRELVSFSSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSLLEDMDGLVDA 1020  
 QY 1021 EBYLVPQGFPCPDPAAGVHHRRHSSSTRSGGDLTGLFESSEERAPSLAPSRG 1080  
 DB 1021 EBYLVPQGFPCPDPAAGVHHRRHSSSTRSGGDLTGLFESSEERAPSLAPSRG 1080  
 QY 1081 AGSDVFDGDLGAKAGLQSLPTPHDSPLOKRSSEDPYPLPSETDGYAVPLTCSQPEVY 1140  
 DB 1081 AGSDVFDGDLGAKAGLQSLPTPHDSPLOKRSSEDPYPLPSETDGYAVPLTCSQPEVY 1140  
 QY 1141 NQPDVRPQPPSPREGPLPARPAGATLBRPKTLSPGKGVVDVAFGAVENPEYLTPQ 1200  
 DB 1141 NQPDVRPQPPSPREGPLPARPAGATLBRPKTLSPGKGVVDVAFGAVENPEYLTPQ 1200  
 QY 1201 GGAAPQHPHPPAFSAFONLYWDDPPERGAPESTFGTPTAENPEYLGLDVPV 1255  
 DB 1201 GGAAPQHPHPPAFSAFONLYWDDPPERGAPESTFGTPTAENPEYLGLDVPV 1255

RESULT 13  
 ADC09593  
 ID ADC09593 standard; protein; 1255 AA.  
 XX  
 AC ADC09593;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Her2/Neu #SEQ ID 594.  
 XX  
 KW Epitope; immunological; vaccine;  
 XX major histocompatibility complex class I; MHC class I; cancer;  
 OS immunisation.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003008537-A2.  
 XX  
 PD 30-JAN-2003.  
 XX  
 PF 29-MAR-2002; 2002WO-US010189.  
 XX  
 PR 06-APR-2001; 2001US-0282211P.  
 PR 07-NOV-2001; 2001US-0337017P.  
 PR 07-MAR-2002; 2002US-0363210P.  
 XX  
 PA (CTL1-) CTL IMMUNOTHERAPIES CORP.  
 XX  
 PI Simard JTL, Diamond DC, Liu L, Xie Z;  
 DR WPI: 2003-248010/24.  
 XX  
 PT Epitope having high affinity for major histocompatibility complex class I  
 PT useful for treating an animal, evaluating immunogenicity of a vaccine or  
 PT therapeutic composition and for diagnosing a disease.  
 XX  
 PS Claim 1; SEQ ID NO 594; 239pp; English.  
 XX  
 CC The invention relates to an isolated epitope polypeptide that has high  
 CC affinity for major histocompatibility complex (MHC) class I, and an  
 CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine  
 CC or immunotherapeutic composition containing an epitope of the invention.  
 CC Compositions of the invention may be used in the treatment of cancer. The  
 CC method can be combined with a radiation therapy, chemotherapy,  
 CC biochemotherapy or surgery. The composition is also useful for evaluating  
 CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC  
 CC -peptide complexes of the invention are useful for determining specific T

cell frequency. This method is useful for evaluating immunological response, by performing the method prior to and subsequent to an immunisation step. Compositions of the invention are useful for diagnosing a disease. The current sequence represents an epitope of the invention with high affinity for MHC class I.

Sequence 1255 AA:

Query Match 100.0%; Score 6815; DB 7; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MELALCMWGLLALLPPGAASVQCTGDMKRLPASPEHLLDMRLYQSCQVQGNL 60
DB 1 MELALCMWGLLALLPPGAASVQCTGDMKRLPASPEHLLDMRLYQSCQVQGNL 60
OY 61 ELTYLPTNASLFLQDIOEVQGYVLAHQVQVPLQRLRYRGVQLFEDNYALAVLNG 120
DB 61 ELTYLPTNASLFLQDIOEVQGYVLAHQVQVPLQRLRYRGVQLFEDNYALAVLNG 120
OY 121 DPLNNTTPTVGTASPGGLRELQRLTEILKGGVLIQRNPOLCYQDTTIIMKQIFKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLTEILKGGVLIQRNPOLCYQDTTIIMKQIFKNNOLA 180
OY 181 LTLIDTNRSRACHPCSPCKSGSRCSSESDCSLTRTVCAAGCARCKGPLEPTDCHEQC 240
DB 181 LTLIDTNRSRACHPCSPCKSGSRCSSESDCSLTRTVCAAGCARCKGPLEPTDCHEQC 240
OY 241 AAGCTGPRHSDCLALHFNHSGICELACPALVTYNTDTPESMPNREGYITGASCVTACP 300
DB 241 AAGCTGPRHSDCLALHFNHSGICELACPALVTYNTDTPESMPNREGYITGASCVTACP 300
OY 301 YNYLSTDVGSCTLYVCPLNHNEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
DB 301 YNYLSTDVGSCTLYVCPLNHNEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
OY 361 IOEFAGCKKIFGSLAFLPESFDGASNTAPLQBPOLQVFETLEBITGYLYISAMPDLP 420
DB 361 IOEFAGCKKIFGSLAFLPESFDGASNTAPLQBPOLQVFETLEBITGYLYISAMPDLP 420
OY 421 DLSVQNTQVTRGRILHNGAYSILTQIGIGISWLGIRSLRELGSALIHNTHLCFVNTV 480
DB 421 DLSVQNTQVTRGRILHNGAYSILTQIGIGISWLGIRSLRELGSALIHNTHLCFVNTV 480
OY 481 PMDOLFRPHOALHTANRPEDECVGEGLACHQLCARHGCMGPGTQCVNCSQFLRGQEC 540
DB 481 PMDOLFRPHOALHTANRPEDECVGEGLACHQLCARHGCMGPGTQCVNCSQFLRGQEC 540
OY 541 VEECRVLQGLPREVYNARHCLPCHPECOFONGSVTCFGBEADQVACAHYKDPFCVARC 600
DB 541 VEECRVLQGLPREVYNARHCLPCHPECOFONGSVTCFGBEADQVACAHYKDPFCVARC 600
OY 601 PSGVPRDLSYMPIMKFPDBEGACQCPINCHSCVDLDDKGPACQARSPLTSTISAVG 660
DB 601 PSGVPRDLSYMPIMKFPDBEGACQCPINCHSCVDLDDKGPACQARSPLTSTISAVG 660
OY 661 ILLVVLGVVFGIILKRRQOKIRKTYMRLLQETELVPLPSGAMPPOAQMRILKETEL 720
DB 661 ILLVVLGVVFGIILKRRQOKIRKTYMRLLQETELVPLPSGAMPPOAQMRILKETEL 720
OY 721 RKVKVLGSGAFGTYKGIWIPDGENVKIPVALIKVRENTSPKANKIIDEAVYMAVGSP 780
DB 721 RKVKVLGSGAFGTYKGIWIPDGENVKIPVALIKVRENTSPKANKIIDEAVYMAVGSP 780
OY 781 YVSRLLGICLTSTVOLVQIMPYGCLLDHVRNRRRLSQDILLNCKQIAKMSYLEVR 840
DB 781 YVSRLLGICLTSTVOLVQIMPYGCLLDHVRNRRRLSQDILLNCKQIAKMSYLEVR 840
OY 841 LVHRDLAARNVAVKSPNHVKTDFGLARLLDIDETEHADGKVPKIMMALESILRRFT 900
DB 841 LVHRDLAARNVAVKSPNHVKTDFGLARLLDIDETEHADGKVPKIMMALESILRRFT 900
OY 901 HQSDVWSYGVTVMEIWTFGAKPYDGI PAKEIPDLLEKGRRLPQPICTIDVYIMVCKWM 960

```

```

DB 901 HQSDVWSYGVTVMEIWTFGAKPYDGI PAKEIPDLLEKGRRLPQPICTIDVYIMVCKWM 960
OY 961 IDSECRPRFRELVEFSRMRADPORFVYIQWEDLGPASPLDSTYRSLLEDMDGDLVDA 1020
DB 961 IDSECRPRFRELVEFSRMRADPORFVYIQWEDLGPASPLDSTYRSLLEDMDGDLVDA 1020
OY 1021 EEYLVPOQGFPCPDPAAGAGVHHRHSSSTRSGGDLTLGLEPSEBARSPALAPSEG 1080
DB 1021 EEYLVPOQGFPCPDPAAGAGVHHRHSSSTRSGGDLTLGLEPSEBARSPALAPSEG 1080
OY 1081 AGSDVFDGDLGMAKGLQSLPTHDPBPLQRYSDDPVYPLPSEFDGYAPLTCSPQPEYV 1140
DB 1081 AGSDVFDGDLGMAKGLQSLPTHDPBPLQRYSDDPVYPLPSEFDGYAPLTCSPQPEYV 1140
OY 1141 NOPVRPOPSPRRGPI PAARPAATLERPYTLSPGNKVQDVAFGAVENPEYLTPO 1200
DB 1141 NOPVRPOPSPRRGPI PAARPAATLERPYTLSPGNKVQDVAFGAVENPEYLTPO 1200
OY 1201 GGAAPQHPPPAFSPAFDNLVYMDQDPPERGAPSTFKGPTJANPEYLGDLVPV 1255
DB 1201 GGAAPQHPPPAFSPAFDNLVYMDQDPPERGAPSTFKGPTJANPEYLGDLVPV 1255

```

RESULT 14  
ADD25484  
ID ADD25484 standard; protein; 1255 AA.

AC ADD25484;  
BT 15-JAN-2004 (first entry)  
XX

Binding domain-immunoglobulin fusion protein-associated protein #19.

XX Binding domain; immunoglobulin; fusion protein; cytosolic;  
XX antidiabetic; immunosuppressive; antidiabetic; antihypertoid;  
XX neuroprotective; hinge region; immunoglobulin heavy chain;  
XX CH2 constant region; CH3 constant region; IgG1;  
XX antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
XX rheumatoid arthritis; myasthenia gravis; Grave's disease;  
XX type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.

OS Unidentified.

PN US2003118592-A1.

PD 26-JUN-2003.

XX 25-JUN-2002; 2002US-00207655.

XX 17-JAN-2001; 2001US-0367358P.

PR 17-JAN-2002; 2002US-00053530.

PR 03-JUN-2002; 2002US-0385691P.

XX (GENE-) GENECRAFT INC.

PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;

XX WPI; 2003-801317/75.

DR New binding domain-immunoglobulin fusion protein, useful for treating a

XX subject having or suspected of having a malignant condition or a B-cell

XX disorder, e.g. melanoma, Grave's disease or autoimmune disease.

XX Disclousure; SEQ ID NO 45; 157bp; English.

XX The invention relates to a binding domain-immunoglobulin fusion protein

XX comprising a binding domain polypeptide that is fused to an

XX immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain

XX CH2 constant region polypeptide that is fused to the hinge region

XX polypeptide, and an immunoglobulin heavy chain CH3 constant region

XX polypeptide that is fused to the CH2 constant region polypeptide. The

CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
 CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
 CC region polypeptide, derived from (a) having 3 or more cysteine residues;  
 CC where the mutated human IgG1 immunoglobulin hinge region polypeptide  
 CC contains 2 cysteine residues, where the first cysteine is not mutated; a  
 CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
 CC (a) having 3 or more cysteine residues, where the mutated human IgG1  
 CC immunoglobulin hinge region polypeptide contains no more than one  
 CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
 CC polypeptide, derived from (a) having 3 or more cysteine residues; where  
 CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
 CC no cysteine residues. The binding domain-immunoglobulin fusion protein is  
 CC capable of at least one immunological activity comprising antibody  
 CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
 CC binding domain polypeptide is capable of specifically binding to an  
 CC antigen. Also included are an isolated polynucleotide encoding the  
 CC binding domain-immunoglobulin fusion protein, a recombinant expression  
 CC construct comprising the polynucleotide (operably linked to a promoter),  
 CC a host cell transformed or transfected with a recombinant expression  
 CC construct, producing the binding domain-immunoglobulin fusion protein, a  
 CC pharmaceutical composition comprising the binding domain-immunoglobulin  
 CC fusion protein or polynucleotide and a carrier, and treating a subject  
 CC having or suspected of having a malignant condition or a B-cell disorder.  
 CC The binding domain-immunoglobulin fusion protein is useful for treating a  
 CC subject having or suspected of having a malignant condition or a B-cell  
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
 CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple  
 CC sclerosis or autoimmune disease. The present sequence is a binding domain  
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
 CC sequence data for this patent formed part of the printed specification  
 CC and is also available in electronic form directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docid=20030118592. The authors have not  
 CC identified the sequences in the printed specification by their SEQ ID  
 CC number therefore none of the sequences can be explicitly identified.

XX Sequence 1255 AA:

CC Query Match 100.0%; Score 6815; DB 7; Length 1255;

CC Query Local Similarity 100.0%; Pred. No. 0; Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELALACWGLLALLPFGAASVCTGTDMKRLPASPETHLDMRLHYGCGVVGNTL 60  
 DB 1 MELALACWGLLALLPFGAASVCTGTDMKRLPASPETHLDMRLHYGCGVVGNTL 60  
 QY 61 EITVLPFNASLSFLDIOEVGVYLIANQVRQVPLQRLIRYRGQLFEDNYALAVLNG 120  
 DB 61 EITVLPFNASLSFLDIOEVGVYLIANQVRQVPLQRLIRYRGQLFEDNYALAVLNG 120  
 QY 121 DPLNNTTPTVGTASPGSLRELQRLSLTEILKGGVLIQRNQLCYDTLLMKDIFHNQDLA 180  
 DB 121 DPLNNTTPTVGTASPGSLRELQRLSLTEILKGGVLIQRNQLCYDTLLMKDIFHNQDLA 180  
 QY 181 LTLIDTNRBRACHPCSPMKGRGSESESDQSLTRTVACAGCARCKGPLETDCHEQC 240  
 DB 181 LTLIDTNRBRACHPCSPMKGRGSESESDQSLTRTVACAGCARCKGPLETDCHEQC 240  
 QY 241 AAGCTGPRKSDCLACIHFHNSGICELHCPALVTYNTDFESMPNREGRTFGASCVTACP 300  
 DB 241 AAGCTGPRKSDCLACIHFHNSGICELHCPALVTYNTDFESMPNREGRTFGASCVTACP 300  
 QY 301 YNYLSTDVGSCTLVGRLNHOEVTABDGTORCKSKPCARVYCYGMEHLRVRVAVTSAN 360  
 DB 301 YNYLSTDVGSCTLVGRLNHOEVTABDGTORCKSKPCARVYCYGMEHLRVRVAVTSAN 360  
 QY 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVETLEETITGLYLSAMPDSL 420  
 DB 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVETLEETITGLYLSAMPDSL 420  
 QY 421 DLSVQNTLOVITGRILHNGAYSLTLOGLSIWLGLRSLRSLGSLALIHNTHLCPVHTV 480  
 DB 421 DLSVQNTLOVITGRILHNGAYSLTLOGLSIWLGLRSLRSLGSLALIHNTHLCPVHTV 480

QY 481 PWDQFLRNPQALHTANRPEDECVBGLACHQLCARGHCWGPRTQCVCNSQFLRGEC 540  
 DB 481 PWDQFLRNPQALHTANRPEDECVBGLACHQLCARGHCWGPRTQCVCNSQFLRGEC 540  
 QY 541 VEECVVLQGLPREYVNAHNCILPCHRCOPQNSVTCFSPBADQCYACAHYKDPFCVARC 600  
 DB 541 VEECVVLQGLPREYVNAHNCILPCHRCOPQNSVTCFSPBADQCYACAHYKDPFCVARC 600  
 QY 601 PSGVPRDLSYMPIMKFPDEBACQPCINCTHSCVDLDDKGPABORASPLTISIAYVG 660  
 DB 601 PSGVPRDLSYMPIMKFPDEBACQPCINCTHSCVDLDDKGPABORASPLTISIAYVG 660  
 QY 661 ILVVVLGVFGLIKRROQIKRYTMRLLQETELVPLTSGAMPNQAQRILKETEL 720  
 DB 661 ILVVVLGVFGLIKRROQIKRYTMRLLQETELVPLTSGAMPNQAQRILKETEL 720  
 QY 721 RKVKVLGSGARFTYKGIWIPDGENVKI PVAIKULRENTSPRANKIIDEAVYAGVGP 780  
 DB 721 RKVKVLGSGARFTYKGIWIPDGENVKI PVAIKULRENTSPRANKIIDEAVYAGVGP 780  
 QY 781 YVSRLLGLCLSTVOLTMPLMYGCLDHYENRGRGLSODLLNMQIAKMSYLEBVR 840  
 DB 781 YVSRLLGLCLSTVOLTMPLMYGCLDHYENRGRGLSODLLNMQIAKMSYLEBVR 840  
 QY 841 LVHRDLAARNVLVKSNNHVKITDFGLARLDDIDETRYHADGKVDIKWMALESILRRPT 900  
 DB 841 LVHRDLAARNVLVKSNNHVKITDFGLARLDDIDETRYHADGKVDIKWMALESILRRPT 900  
 QY 901 HOSDWSYGVTYWELMTGAKPYDGI PARBIPLDLKESRLPQPICTIDVYMIWVKCM 960  
 DB 901 HOSDWSYGVTYWELMTGAKPYDGI PARBIPLDLKESRLPQPICTIDVYMIWVKCM 960  
 QY 961 IDSECRPFRELVSFESMARDPORFVYI QNEDLGPASPLDSTFPRSLLDDMDGLVDA 1020  
 DB 961 IDSECRPFRELVSFESMARDPORFVYI QNEDLGPASPLDSTFPRSLLDDMDGLVDA 1020  
 QY 1021 BEYLVPOGFFCPDPAPAGGVVHRHRSJSTRSGGDLTLGLAEPSEBAPRSLAPSEB 1080  
 DB 1021 BEYLVPOGFFCPDPAPAGGVVHRHRSJSTRSGGDLTLGLAEPSEBAPRSLAPSEB 1080  
 QY 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
 DB 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
 QY 1141 NOPDVRPQPSRBSGLPAARPAATLRLPKTSLSGKGVVADVAPFGAVENPEYLTPQ 1200  
 DB 1141 NOPDVRPQPSRBSGLPAARPAATLRLPKTSLSGKGVVADVAPFGAVENPEYLTPQ 1200  
 QY 1201 GGAAPQPPPPAFSPAFNLYWDDPBERGAPSTFKGTPTAENPEYLGLDV 1255  
 DB 1201 GGAAPQPPPPAFSPAFNLYWDDPBERGAPSTFKGTPTAENPEYLGLDV 1255

RESULT 15  
 ID ADE63281  
 ID ADE63281 standard, protein, 1255 AA.  
 XX ADE63281;  
 DT 29-JAN-2004 (first entry)  
 DE Human Protein P04626, SEQ ID NO 9218.  
 XX Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 OS spared nerve injury; SN1; Chung.  
 XX Homo sapiens.  
 PN MO2003016475-A2.  
 XX 27-FEB-2003.  
 XX

PF 14-AUG-2002; 2002WO-US025765.  
XX  
XX 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
PI  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
XX WPI, 2003-268312/26.  
DR GENBANK; P04626.  
PT  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.  
XX  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNR)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 1255 AA;  
Query Match 100.0%; Score 6815; DB 7; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 YNVLSTVSGCTLVCPILHNOETVAEDGTORCEKSKRCARVCYGLGMEHLREPAVTSAN 360  
DB 301 YNVLSTVSGCTLVCPILHNOETVAEDGTORCEKSKRCARVCYGLGMEHLREPAVTSAN 360  
QY 361 IOEFAGCKKIFGSLAFIPESFDGDPASNTAPLOEEOLOVFETLEITNGYLYISAMPBSLP 420  
DB 361 IOEFAGCKKIFGSLAFIPESFDGDPASNTAPLOEEOLOVFETLEITNGYLYISAMPBSLP 420  
QY 421 DLSVFONLQVLRGRILHNGAVSLTLQGLISWGLRSLRBLGSGIALIHNTHLCPVTV 480  
DB 421 DLSVFONLQVLRGRILHNGAVSLTLQGLISWGLRSLRBLGSGIALIHNTHLCPVTV 480  
QY 481 FMDOLFNNPHOALHTANRPEDECVGSGLACHQICANGHCMBPPTQCVNCSQLRQEC 540  
DB 481 FMDOLFNNPHOALHTANRPEDECVGSGLACHQICANGHCMBPPTQCVNCSQLRQEC 540  
QY 541 VEECRVLOGLPREYVNAHCLPCHPEGOPONGSVTCGPEADQCVACHYKDPFCVARC 600  
DB 541 VEECRVLOGLPREYVNAHCLPCHPEGOPONGSVTCGPEADQCVACHYKDPFCVARC 600  
QY 601 PSGVKPDLSTYMPIMKFPDEBGACQPCPINCTHSCVDLDDKGCAPAEORASPLTISIAYVG 660  
DB 601 PSGVKPDLSTYMPIMKFPDEBGACQPCPINCTHSCVDLDDKGCAPAEORASPLTISIAYVG 660  
QY 661 ILLVVLGVVFGIILIKRQOKIRKRYTMRRLIOETELVBLTPSGAMPNOAQRILKTEL 720  
DB 661 ILLVVLGVVFGIILIKRQOKIRKRYTMRRLIOETELVBLTPSGAMPNOAQRILKTEL 720  
QY 721 RKVNVLSGANGVYKGIWIPDGENVKIPVAIKYLRNTSPKAKKEILDEAYVWAGVSP 780  
DB 721 RKVNVLSGANGVYKGIWIPDGENVKIPVAIKYLRNTSPKAKKEILDEAYVWAGVSP 780  
QY 781 YVSRLLGICLTSTVQLVTLQMLPYGCLLDHVENRGRISODLLMWQIAKMSLYEDVR 840  
DB 781 YVSRLLGICLTSTVQLVTLQMLPYGCLLDHVENRGRISODLLMWQIAKMSLYEDVR 840  
QY 841 LVHBDLAARNVLVKSPNVKTTDGLARLDDIDETEHADGKPIKMMALSLIRRF 900  
DB 841 LVHBDLAARNVLVKSPNVKTTDGLARLDDIDETEHADGKPIKMMALSLIRRF 900  
QY 901 HQSVWSGYVTVMELMTFGAKPYDGIPIAREIPDLLEKGERLPOPPICITIDVYMIWVKCM 960  
DB 901 HQSVWSGYVTVMELMTFGAKPYDGIPIAREIPDLLEKGERLPOPPICITIDVYMIWVKCM 960  
QY 961 IDSECRPRFRELVEFSRMAADPQRFVVTQNEDELGPASPLDSTFYRSLLEDMDGLVDA 1020  
DB 961 IDSECRPRFRELVEFSRMAADPQRFVVTQNEDELGPASPLDSTFYRSLLEDMDGLVDA 1020  
QY 1021 EBYLVPOQGFCDPPAPGAGGMVHHRSSSTRSGGDLTLGLPSEBRARSPLABSEG 1080  
DB 1021 EBYLVPOQGFCDPPAPGAGGMVHHRSSSTRSGGDLTLGLPSEBRARSPLABSEG 1080  
QY 1081 AGSDVPDGLGMSGAKGLOSLPTHDPSPLOKYSDDPTVPLPSERDGYVAPLTCSQPPRYV 1140  
DB 1081 AGSDVPDGLGMSGAKGLOSLPTHDPSPLOKYSDDPTVPLPSERDGYVAPLTCSQPPRYV 1140  
QY 1141 NQPDVRPQPSPREGPLPAAPAGATLERPKTLSPKNGVVKDVFARGAIVENBEYLTPO 1200  
DB 1141 NQPDVRPQPSPREGPLPAAPAGATLERPKTLSPKNGVVKDVFARGAIVENBEYLTPO 1200  
QY 1201 GGAAPQPHPPAPSPADNLYWQDPPERGAAPSTFKGTPTAENPEYLGIDVY 1255  
DB 1201 GGAAPQPHPPAPSPADNLYWQDPPERGAAPSTFKGTPTAENPEYLGIDVY 1255

Search completed: January 3, 2006, 11:12:19  
Job time : 204 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using ew model

Run on: January 3, 2006, 11:05:47 ; Search time 50 Seconds  
(without alignments)  
2415.041 Million cell updates/sec

Title: US-09-930-125-2  
Perfect score: 6815  
Sequence: 1 METALCRMGLLLALLPRCA.....TFKGTPTANPEYGLDVPV 1255

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	1 A24571	protein-tyrosine k
2	5997	88.0	1260	1 TVRTNU	protein-tyrosine k
3	5993.5	87.9	1254	2 I48161	p-185 precursor
4	3169	46.5	1210	1 GQHUE	epidermal growth f
5	3145	46.1	1210	2 A53183	epidermal growth f
6	3122.5	45.8	1223	1 TVGHIV	epidermal growth f
7	3004.5	44.1	1308	2 A47253	epidermal growth f
8	2708	39.7	1166	1 S06142	protein-tyrosine k
9	2432.5	35.7	1342	2 A36223	kinase-related tra
10	2347.5	34.4	1339	2 JCA387	epidermal growth f
11	1765.5	25.9	698	1 TVFVLV	protein-tyrosine k
12	1702	25.0	604	1 TVYVUH	protein-tyrosine k
13	1653.5	24.3	1330	1 GQFPE	epidermal growth f
14	1645	24.1	544	2 S35745	protein-tyrosine k
15	1638	24.0	545	2 S00727	kinase-related tra
16	1621	23.8	540	2 B44776	protein-tyrosine k
17	1619	23.8	540	1 TVFVEB	protein-tyrosine k
18	1536	22.5	644	2 A36325	epidermal growth f
19	1301	19.1	1323	2 E88257	protein-tyrosine k
20	1301	19.1	1374	2 S70712	protein-tyrosine k
21	1211	17.8	1369	2 S70713	protein-tyrosine k
22	1177	17.3	1717	1 A45558	epidermal growth f
23	1155	16.9	527	2 A42032	epidermal growth f
24	997.5	14.6	843	2 A27131	epidermal growth f
25	814.5	12.0	346	2 S13807	protein-tyrosine k
26	757	11.1	311	2 S13808	protein-tyrosine k
27	736	10.8	1363	2 T43220	insulin-like growth
28	717	10.5	1382	1 INHUR	insulin receptor p
29	710	10.4	1383	2 A36080	insulin receptor p

30	709.5	10.4	1372	2 A34157	insulin receptor p
31	703.5	10.3	1607	2 T43212	insulin-like growth
32	693.5	10.2	1300	2 A36502	insulin receptor-r
33	682.5	10.0	1477	2 T18534	protein-tyrosine k
34	681	10.0	1268	2 B36502	insulin receptor-r
35	651	9.6	1367	1 IGHU1	insulin-like growth
36	642	9.4	1371	2 A33837	insulin-like growth
37	627	9.2	1390	2 T30346	insulin receptor -
38	624.5	9.2	2148	1 A56081	insulin receptor
39	622	9.1	2101	2 S57245	insulin receptor (
40	607	8.9	987	2 A54092	protein-tyrosine k
41	591	8.7	1091	2 S33596	protein-tyrosine k
42	590.5	8.7	952	2 I50612	protein-tyrosine k
43	589.5	8.7	977	2 S49004	tyrosine kinase Mp
44	587	8.6	1114	1 S05582	protein-tyrosine k
45	585	8.6	987	2 I48652	mouse developmenta

## ALIGNMENTS

RESULT 1  
A24571  
protein-tyrosine kinase (BC 2.7.1.112) erbB2 precursor - human  
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erbB  
C:Species: Homo sapiens (man)  
C>Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 05-Oct-2004  
C:Accession: A24571; A25491; A44488; B44188; I59509; I57622  
R:Yamamoto, T.; Ikawa, S.; Miyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;  
Nature 319, 230-234, 1996  
A>Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth f  
A:Reference number: A24571; MUID:86118663; PMID:3003577  
A:Accession: A24571  
A:Molecule type: mRNA  
A:Residues: 1-1255 <YAM>  
A:Cross-references: UNIPROT:P04626; UNIPARC:UPI000003F55F; GB:X03363; NID:g31197; PIDN:C  
R:Semba, K.; Yamamoto, N.; Toyoshima, K.; Yamamoto, T.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985  
A>Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epiderm  
A:Reference number: A25491; MUID:86016729; PMID:2995967  
A:Accession: A25491  
A:Molecule type: DNA  
A:Residues: 737-1031 <SEM>  
A:Cross-references: UNIPARC:UPI000016A0A7; GB:M11767; NID:G182163; PIDN:AAA35808.1; PID:  
R:Cousens, L.; Yang-Peng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P  
Science 230, 1132-1139, 1985  
A>Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromos  
A:Reference number: A44188; MUID:86070181; PMID:2999974  
A:Accession: A44188  
A:Molecule type: DNA  
A:Residues: 740-910 <COU>  
A:Cross-references: UNIPARC:UPI000016AA26; GB:M12036; NID:G183988; PIDN:AAA35978.1; PID:  
A:Accession: B44188  
A:Molecule type: mRNA  
A:Residues: 1-517, 'RALU', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>  
A:Cross-references: UNIPARC:UPI00001725C7; GB:M11730; NID:G183986  
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.  
Science 229, 974-976, 1985  
A>Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.  
A:Reference number: I59509; MUID:85272597; PMID:2992089  
A:Accession: I59509  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 832-909 <REX>  
A:Cross-references: UNIPARC:UPI0000070A3F; GB:L29395; NID:G459807; PIDN:AAA35809.1; PID:  
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.  
Mol. Cell. Biol. 7, 2597-2601, 1987  
A>Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional  
A:Reference number: I57622; MUID:87286698; PMID:3033551  
A:Accession: I57622  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-191 <TAU>



A:Cross-references: UNIPARC:UP1000000427; GB:M6792; NID:9183983; PIDN:AA58637.1; PID:  
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30  
 C:Genetics:  
 A:Gene: GDB:ERBB2; NGI; NEU; HER-2  
 A:Cross-references: GDB:120613; OMIM:164870  
 A:Map position: 17q21.1-17q21.1  
 A:Introns: 25/1; 75/3; 147/1; 863/3  
 A>Note: the list of introns is incomplete  
 C:Function:  
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
 kinase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:122-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>  
 F:22-653/Domain: extracellular #status predicted <EXT>  
 F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>  
 F:139-605/Domain: EGF receptor extracellular domain repeat <EE2>  
 F:1654-675/Domain: transmembrane #status predicted <TM>  
 F:676-1255/Domain: intracellular #status predicted <INT>  
 F:718-983/Domain: protein kinase homology <KIN>  
 F:726-734/Region: protein kinase ATP-binding motif  
 F:68,124,181,259,330,571,629/Binding site: carbohydrate (Asn) (covalent) #status predict  
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F:753/Active site: Lys #status predicted  
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 100.0%; Score 6815; DB 1; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-277;  
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELALCWMGLLALLPFGASTVQCTGDMKRLPASPEHLLDMRLHYOCQVQGNL 60  
 DB 1 MELALCWMGLLALLPFGASTVQCTGDMKRLPASPEHLLDMRLHYOCQVQGNL 60  
 QY 61 ELYLPPTNASLSFLDIOBQGVYLIANQVRQVPLQRLRYRGTLFEDNYALAVLDNG 120  
 DB 61 ELYLPPTNASLSFLDIOBQGVYLIANQVRQVPLQRLRYRGTLFEDNYALAVLDNG 120  
 QY 121 DPLNNTTAVTASPGSLRELOSLTEILKGGVLIQRNPOLCYQTLIMKDIFFHNNOLA 180  
 DB 121 DPLNNTTAVTASPGSLRELOSLTEILKGGVLIQRNPOLCYQTLIMKDIFFHNNOLA 180  
 QY 181 LTLIDTNBSRACHPCSPCKSGRCWSESEDCQSLTRTVACAGGACRCKGPLPTDCHEQC 240  
 DB 181 LTLIDTNBSRACHPCSPCKSGRCWSESEDCQSLTRTVACAGGACRCKGPLPTDCHEQC 240  
 QY 241 AAGCTGPRHSDCLAHFNHSGICELHCPALVTYNTTFESMPNDEGRYTFGASCVTACP 300  
 DB 241 AAGCTGPRHSDCLAHFNHSGICELHCPALVTYNTTFESMPNDEGRYTFGASCVTACP 300  
 QY 301 YNYLSTDVSGCTLVPLRNOEVTADGQRCCKSKPCARVYGYGMHLEVRVAVTSAN 360  
 DB 301 YNYLSTDVSGCTLVPLRNOEVTADGQRCCKSKPCARVYGYGMHLEVRVAVTSAN 360  
 QY 361 IOEFAGCKKIFGSLAFLESPFDGASNTAPLQPEQLQVFTLEETITGYLISAMPDILP 420  
 DB 361 IOEFAGCKKIFGSLAFLESPFDGASNTAPLQPEQLQVFTLEETITGYLISAMPDILP 420  
 QY 421 DLSVFQNTQVIRGRILHNGAVSLTQIGIGISWLGIRSLRELSGIALIHHNTHLCFVHTV 480  
 DB 421 DLSVFQNTQVIRGRILHNGAVSLTQIGIGISWLGIRSLRELSGIALIHHNTHLCFVHTV 480  
 QY 481 PMDOLFRPHOALHTANRBECEVGEGLACHQLCARHGCMGPPTOCVNCQSPILRGEC 540  
 DB 481 PMDOLFRPHOALHTANRBECEVGEGLACHQLCARHGCMGPPTOCVNCQSPILRGEC 540  
 QY 541 VEECEVLQGLPREYVNAHCLCPHEPCOPONGSVTCFGEADQCVACAHYKDPFCVARC 600  
 DB 541 VEECEVLQGLPREYVNAHCLCPHEPCOPONGSVTCFGEADQCVACAHYKDPFCVARC 600  
 QY 601 PSGVPRDLSYMPIMKPFDEEGACQPCPINCTHSCVDLDDKGPAPQASPLTISAVVG 660  
 DB 601 PSGVPRDLSYMPIMKPFDEEGACQPCPINCTHSCVDLDDKGPAPQASPLTISAVVG 660

DB 601 PSGVPRDLSYMPIMKPFDEEGACQPCPINCTHSCVDLDDKGPAPQASPLTISAVVG 660  
 QY 661 ILVVLGVVGGILIKRQOKIRKYTWRRLLQETELVEPLTPSGAMPNOAQMRILKETEL 720  
 DB 661 ILVVLGVVGGILIKRQOKIRKYTWRRLLQETELVEPLTPSGAMPNOAQMRILKETEL 720  
 QY 721 RKVVLGSGAFGVYKGIWIPDGENVKI PVAIKYLRENTSFKANKEIIDEAYVMAVGSP 780  
 DB 721 RKVVLGSGAFGVYKGIWIPDGENVKI PVAIKYLRENTSFKANKEIIDEAYVMAVGSP 780  
 QY 781 VYSLRLGICLTSTVQVLTQMPVGCILDHYENNGRIGSODLMMQOIAAGMSYLEDVR 840  
 DB 781 VYSLRLGICLTSTVQVLTQMPVGCILDHYENNGRIGSODLMMQOIAAGMSYLEDVR 840  
 QY 841 LVHDDLAAARNVLKSPNNVKTIDFGRLRLDIDETEVHADGKVPKIMMALESILRRFT 900  
 DB 841 LVHDDLAAARNVLKSPNNVKTIDFGRLRLDIDETEVHADGKVPKIMMALESILRRFT 900  
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMTMVKCM 960  
 DB 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMTMVKCM 960  
 QY 961 IDSECRPRFRELVSFBRMADPQRFVLTQEDIGPASPLDSTYRSILBEDMDGLVDA 1020  
 DB 961 IDSECRPRFRELVSFBRMADPQRFVLTQEDIGPASPLDSTYRSILBEDMDGLVDA 1020  
 QY 1021 EBYLVPOQGFPCPDPAAGAGVHRRSSSTRSGGDLTLGLPSEBEARSPPLAPSEG 1080  
 DB 1021 EBYLVPOQGFPCPDPAAGAGVHRRSSSTRSGGDLTLGLPSEBEARSPPLAPSEG 1080  
 QY 1081 AGSDVFDGLMGAKGLQSLPTHDPBLQRYSDPVPYLPSEFDGYAPLTCSPQPEYV 1140  
 DB 1081 AGSDVFDGLMGAKGLQSLPTHDPBLQRYSDPVPYLPSEFDGYAPLTCSPQPEYV 1140  
 QY 1141 NOPVRPOPSPBRGCPPLAARPAATLBRPKTSLPGKNGVQDVAFGGAENBEYLTPQ 1200  
 DB 1141 NOPVRPOPSPBRGCPPLAARPAATLBRPKTSLPGKNGVQDVAFGGAENBEYLTPQ 1200  
 QY 1201 GGAAPQHPPPAFPAFDNLTYMDQDPPERGAPSTFKGPTAANPEYLGIDVY 1255  
 DB 1201 GGAAPQHPPPAFPAFDNLTYMDQDPPERGAPSTFKGPTAANPEYLGIDVY 1255

## RESULT 2

TYRINU

protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 05-Oct-2004

C:Accession: A24562; A61204

R:Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.

Nature 319, 226-230, 1986

A>Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.

A:Reference number: A24562; MUID:86118662; PMID:3945311

A:Accession: A24562

A:Molecule type: mRNA

A:Residues: 1-1260 <BAR>

A:Cross-references: UNIPROT:P06494; UNIPARC:UP10000161B83; EMBL:X03362; NID:956745; PIDN:

R:Maui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,

Carcinogenesis 12, 1975-1978, 1991

A>Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no m

2-thiazolyl]formamide or N-methyl-N-nitrosourea.

A:Reference number: A61204; MUID:92035293; PMID:1682063

A:Accession: A61204

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 637-663, 'V', 665-702 <MS>

A:Cross-references: UNIPARC:UP100001725C8

A>Note: authors translated the codon GCA for residue 25 as Val

C:Genetics:

A:Gene: neu  
 C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
 F:1-19/Domain: signal sequence #status predicted <SIG>



F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>  
 F:658-980/Domain: transmembrane #status predicted <TMN>  
 F:723-988/Domain: protein kinase homology <KIN>  
 F:731-739/Region: protein kinase ATP-binding motif  
 F:71,191,263,535,576,634/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted  
 F:758/Active site: Lys #status predicted  
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 88.0%; Score 5997; DB 1; Length 1260;  
 Best Local Similarity 87.9%; Pred. No. 4,3e-243;  
 Matches 1105; Conservative 49; Mismatches 101; Indels 2; Gaps 2;

```

QY 1 MELAALCRWGLLALLPFGAASVQVCTGDMKRLRPAPEETHLDMRLHYOGCQVVOGNTL 60
DB 4 MELAAWCMKRGFLALLPFGIAGTQVCTGDMKRLRPAPEETHLDMRLHYOGCQVVOGNTL 63
QY 61 ELTYLPTNASLSFLQDIDEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
DB 64 ELTYVPANASLSFLQDIDEVQGYVLIANQVQVPLQRLIRVGTQLFEDKXALAVLDNR 123
QY 121 DELANTTPVT-GASPEGLELQRLSTLILKGVLIQRNPQLCYQDTILMKDIFHKNOQL 179
DB 124 DQDNVAASPTGRTPEGLRELQRLSTLILKGVLIQRNPQLCYQDMVLKQVFRKNQQL 183
QY 180 ALTLIDTNSRACHPCSPMKGSRGESSBDCQSLTRTVACAGCARCGPLPTDCCHQ 239
DB 184 AAVDIDTNSRACHPCSPARCKDNKCGESBDCQILITGITCSGCRCKGRPLTDCCHQ 243
QY 240 CAAGCTGPRKSDCLACLFHNHSGICELACPALVTYNTDFESMPNDEGRYTFGASCVTAC 299
DB 244 CAAGCTGPRKSDCLACLFHNHSGICELHCPALVTYNTDFESMHNDEGRYTFGASCVTTC 303
QY 300 PNNYSTDVGSCTLVCPRLNQVYTAADGTQRCCKSKPCARVCYGIAMHLEAVRAVTA 359
DB 304 PNNYSTDVGSCTLVCPRLNQVYTAADGTQRCCKSKPCARVCYGIAMHLEAVRAVTS 363
QY 360 NIOEPAGCKKIFGSLAFLPESFGDPASNTAPLQPOLQVFETLSEITGYLISAMPDSL 419
DB 364 NIOEPAGCKKIFGSLAFLPESFGDPASNTAPLQPOLQVFETLSEITGYLISAMPDSL 423
QY 420 PDLVSFQNLQVIRGRILHNHGAVALTQGLGISMVGLSRSLRELSGLALIHNTHLCEVNT 479
DB 424 RLSVSFQNLQVIRGRILHNHGAVALTQGLGISMVGLSRSLRELSGLALIHNTHLCEVNT 483
QY 480 VPMWQQLFRPHQALHTANRPEDE-CVGEGLACHQICAGRCCKGPRPTCVNCSQPLRQ 538
DB 484 VPMWQQLFRPHQALHTANRPEDE-CVGEGLACHQICAGRCCKGPRPTCVNCSQPLRQ 543
QY 539 ECVVEGRVLTQGLPREVYNARHCLPCHREGOPONGSVTCGPEADOCVAHAHKDPPCYA 598
DB 544 ECVVEGRVLTQGLPREVYNARHCLPCHREGOPONGSVTCGPEADOCVAHAHKDPPCYA 603
QY 599 RCPGSGVPLDLYMPPIWKPFDEBEGACQPCINCTHSCVDLDDKCPAEQASPLTISIAY 658
DB 604 RCPGSGVPLDLYMPPIWKPFDEBEGACQPCINCTHSCVDLDDKCPAEQASPLTISIAY 663
QY 659 VGILLVVLGVVFGIILKRRQKIRKTYMRLLQETELVEPLTPSGAMNQOMRLKST 718
DB 664 VGILLVVLGVVFGIILKRRQKIRKTYMRLLQETELVEPLTPSGAMNQOMRLKST 723
QY 719 ELRKVAVLGSAGFGTYVKGIMI PDGENVYKI PVAIKYLRNTSPKANKELLDAYVMAAGV 778
DB 724 ELRKVAVLGSAGFGTYVKGIMI PDGENVYKI PVAIKYLRNTSPKANKELLDAYVMAAGV 783
QY 779 SPYVSLRLGICLTSTVQVLTQVLMPIYGCLLDHYRENRGLSQDGLNMVCMQIAKMSYLED 838
DB 784 SPYVSLRLGICLTSTVQVLTQVLMPIYGCLLDHYRENRGLSQDGLNMVCMQIAKMSYLED 843
QY 839 VALVHRDLAARVNLVKSPMHVKITDGLARLLDIDETRYHADGKRPITPMALLESILRR 898
DB 844 VALVHRDLAARVNLVKSPMHVKITDGLARLLDIDETRYHADGKRPITPMALLESILRR 903

```

QY 899 FTHQSDVMSYGVTVWELMTFGAKPYDGIAPAREIPDLLEKGERLPOPPICITIDVYIMVVC 958  
 DB 904 FTHQSDVMSYGVTVWELMTFGAKPYDGIAPAREIPDLLEKGERLPOPPICITIDVYIMVVC 963  
 QY 959 WMIDECPRPRFELVSEFSRNARDQRRVYVIONELGRPASPLDSFTFYSLLEDMDMDGLV 1018  
 DB 964 WMIDECPRPRFELVSEFSRNARDQRRVYVIONELGRPASPLDSFTFYSLLEDMDMDGLV 1023  
 QY 1019 DAEYLVPOQGFPCDPPAGAGWVHHRRSSSTSGGDDLTLGLPSEBEAPRSPLAPS 1078  
 DB 1024 DAEYLVPOQGFPCDPPAGAGWVHHRRSSSTSGGDDLTLGLPSEBEAPRSPLAPS 1083  
 QY 1079 EGAGSDVDFDGLGMAAGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVVAPLTCSPQPE 1138  
 DB 1084 EGAGSDVDFDGLGMAAGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVVAPLTCSPQPE 1143  
 QY 1139 YVNOQDVPRPQPSPEEGPLPAARPRGATLIERKXTLSPGKNGVYKQVAPGAVENPEYLT 1198  
 DB 1144 YVNOQDVPRPQPSPEEGPLPAARPRGATLIERKXTLSPGKNGVYKQVAPGAVENPEYLT 1203  
 QY 1199 POGGAPOPHPPAPSPAFNDLYYWDQDPPERGAPSTFKGTPTAENPEYLGADV 1255  
 DB 1204 POGGAPOPHPPAPSPAFNDLYYWDQDPPERGAPSTFKGTPTAENPEYLGADV 1260

RESULT 3  
 148161  
 p-185 precursor - golden hamster  
 C/Dates: Mesocricetus auratus (golden hamster)  
 C/Species: Mesocricetus auratus (golden hamster)  
 C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Oct-2004  
 C/Accession: 148161  
 R/Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa, Gene 140, 251-255, 1994  
 A/Title: Cloning and activation of the Syrian hamster neu proto-oncogene.  
 A/Reference number: 148161; M01D:94193007; PMID:7908275  
 A/Accession: 148161  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-1254 <RES>  
 A/Cross-references: UNIPROT:060553; UNIPARC:UP1000012A111; GB:D16295; NID:g493236; P1DN: A/Gene: neu  
 C/Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology  
 C/Keywords: ATP  
 F:718-983/Domain: protein kinase homology <KIN>  
 F:726-734/Region: protein kinase ATP-binding motif

Query Match 87.9%; Score 5993.5; DB 2; Length 1254;  
 Best Local Similarity 87.7%; Pred. No. 5.9e-243;  
 Matches 1101; Conservative 57; Mismatches 96; Indels 1; Gaps 1;

```

QY 1 MELAALCRWGLLALLPFGAASVQVCTGDMKRLRPAPEETHLDMRLHYOGCQVVOGNTL 60
DB 1 MELAAWCMKRGFLALLPFGIAGTQVCTGDMKRLRPAPEETHLDMRLHYOGCQVVOGNTL 60
QY 61 ELTYLPTNASLSFLQDIDEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
DB 64 ELTYVPANASLSFLQDIDEVQGYVLIANQVQVPLQRLIRVGTQLFEDKXALAVLDNR 120
QY 121 DELANTTPVTGASPEGLELQRLSTLILKGVLIQRNPQLCYQDTILMKDIFHKNOQL 180
DB 124 DPLDNTVATGRTPEGLRELQRLSTLILKGVLIQRNPQLCYQDTILMKDIFHKNOQL 180
QY 181 LTLIDTNSRACHPCSPMKGSRGESSBDCQSLTRTVACAGCARCGPLPTDCCHQ 240
DB 184 LTLIDTNSRACHPCSPMKGSRGESSBDCQSLTRTVACAGCARCGPLPTDCCHQ 240
QY 241 AAGCTGPRKSDCLACLFHNHSGICELACPALVTYNTDFESMPNDEGRYTFGASCVTAC 300
DB 244 AAGCTGPRKSDCLACLFHNHSGICELHCPALVTYNTDFESMHNDEGRYTFGASCVTTC 300
QY 301 PNNYSTDVGSCTLVCPRLNQVYTAADGTQRCCKSKPCARVCYGIAMHLEAVRAVTSAN 360

```

Db 301 YNVLSTEWGSCITVCPINNQEVTAEDGTQRCERKSCSKARVCYGLGMEHLRGARATISAN 360  
 QY 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLQPPOLQVETLEETITGLYISAMPDILP 420  
 Db 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLQPPOLQVETLEETITGLYISAMPDILH 420  
 QY 421 DLVSFQNTQVIRGRILHNGAVSLTLQGLGISWLGSLRLSELGSLALIHNTLHCFVHTV 480  
 Db 421 DLVSFQNTQVIRGRILHNGAVSLTLQGLGISWLGSLRLSELGSLALIHNTLHCFVHTV 480  
 QY 461 PMDQFLRPHQALHTTARPREDECEGEGLAGHQLCARHGCMWPGTQVCNCGEQRLRGEC 540  
 Db 461 PMDQFLRPHQALHTTARPREDECEGEGLAGHQLCARHGCMWPGTQVCNCGEQRLRGEC 540  
 QY 541 VEECHVLQGLPREYVNAHRCLECHPECOPONGSVTCFGEADQCAACHYKDPFCVARC 600  
 Db 541 VEECHVLQGLPREYVNAHRCLECHPECOPONGSVTCFGEADQCAACHYKDPFCVARC 600  
 QY 541 VKECHVWKGFLPREYVNGKCLFCHPECOPONSTETCTGSEADQCAACHYKDPFCVARC 600  
 Db 601 PSQVPRDLSTYMPIMKYPDEEGWCOPCPIINCHSCVDLDRGCPAQRASPATSIATVVG 660  
 QY 661 ILVVVLGVVPGIILKKRQOKIRKTYTMRLLQETLVEPLTPSGAMPNOAQMRILKETEL 720  
 Db 661 ILVFLVIGVVGILLKRRQOKIRKTYTMRLLQETLVEPLTPSGAMPNOAQMRILKETEL 720  
 QY 721 RKVKVLGSGAFGTVYKGIWIPDGEVVKIPVAIKVIRENTSPKANKEILDEAVYMGVSGP 780  
 Db 721 RKVKVLGSGAFGTVYKGIWIPDGEVVKIPVAIKVIRENTSPKANKEILDEAVYMGVSGP 780  
 QY 781 YVSRLLGICITSTVQLMPYGLLDHVRNRRGLSDILLNMCQIANKMSYLEDVR 840  
 Db 781 YVSRLLGICITSTVQLMPYGLLDHVRNRRGLSDILLNMCQIANKMSYLEDVR 840  
 QY 841 LVHRLAARNLVKSPNHVKITDTEFGALRLDIDETEHADGKVPKIMTALESILIRRF 900  
 Db 841 LVHRLAARNLVKSPNHVKITDTEFGALRLDIDETEHADGKVPKIMTALESILIRRF 900  
 QY 901 HOSDWWSGVTVWELMTFGAKPYDGI PAKEIPDLLEKGERLPPPICTIDVYMIWVKCM 960  
 Db 901 HOSDWWSGVTVWELMTFGAKPYDGI PAKEIPDLLEKGERLPPPICTIDVYMIWVKCM 960  
 QY 961 IDSECRPRRELVSFSSRMARDPQRFVVIQNEIDLCPASPLDSTFYRSLLEDMDGLVDA 1020  
 Db 961 IDSECRPRRELVSFSSRMARDPQRFVVIQNEIDLCPASPLDSTFYRSLLEDMDGLVDA 1020  
 QY 1021 BEYLVPOQGFCEPDPAFGAGVHHRHSSSTRSGGDLTGLBSEBEAPRSLABSEG 1080  
 Db 1021 BEYLVPOQGFCEPDPAFGAGVHHRHSSSTRSGGDLTGLBSEBEAPRSLABSEG 1080  
 QY 1081 AGSDVFDGDLGMAKGLQSLPTHPSPLOQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
 Db 1081 AGSDVFDGDLGMAKGLQSLPTHPSPLOQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
 QY 1141 NQPDVROPSPREBPLPAARAGATLBRPKTLSPGKGVVQVDAFGAVENPEYLTPO 1200  
 Db 1141 NQPDVROPSPREBPLPAARAGATLBRPKTLSPGKGVVQVDAFGAVENPEYLTPO 1200  
 QY 1201 GGSAQOPH-PPALCPAFDMLYWDQDPSERGSPPMTFECTPAENPEYLTGLDVPV 1254  
 Db 1201 GGSAQOPH-PPALCPAFDMLYWDQDPSERGSPPMTFECTPAENPEYLTGLDVPV 1254

## RESULT 4

GQHUB  
 epidermal growth factor receptor precursor - human  
 N>Contactin: protein-tyrosine kinase (EC 2.7.1.112) erbb  
 C:Species: Homo sapiens (man)  
 C:Date: 15-Nov-1984 #sequence revision 27-Nov-1985 #text change 05-Oct-2004  
 C:Accession: A00641; A25772; S30024; A30642; A43615; A23062; A05281; A60143; A33  
 R:Ulrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Y  
 Ig, P.H.  
 Nature 309, 418-425, 1984

A>Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of  
 A:Reference number: A00641; MUID:84219729; PMID:6328312  
 A:Accession: A00641  
 A:Molecule type: mRNA  
 A:Residues: 1-1210 <URL>  
 A:Cross-references: UNIPROT:P00533; UNIPARC:UPI0000050F30; EMBL:X00588; NID:g31113; PIDN  
 A>Note: the authors translated the codon AAG for residue 540 as Asn  
 R:Rishi, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985  
 A>Title: Characterization and sequence of the promoter region of the human epidermal gro  
 A:Reference number: A25772; MUID:85270438; PMID:2991899  
 A:Accession: A25772  
 A>Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-29 <ISH>  
 A:Cross-references: UNIPARC:UPI000016A882; GB:M1234; NID:g181981; PIDN:AAA2370.1; PID:  
 R:Haley, J.; Whittle, N.; Bennett, P.; Knechtgen, D.; Ulrich, A.; Waterfield, M.  
 Oncogene Res. 1, 375-396, 1987  
 A>Title: The human EGF receptor gene: structure of the 110 kb locus and identification o  
 A:Reference number: S30024; MUID:88217333; PMID:3329716  
 A:Accession: S30024  
 A:Molecule type: DNA  
 A:Residues: 1-29 <HA2>  
 A:Cross-references: UNIPARC:UPI000016A882; EMBL:X06370; NID:g31118; PIDN:CAA29668.1; PID  
 R:Haley, J.D.; Waterfield, M.D.  
 J. Biol. Chem. 266, 1746-1753, 1991  
 A>Title: Contributory effects of de Novo transcription and premature transcript terminat  
 A:Reference number: A38672; MUID:91107677; PMID:1988448  
 A:Accession: A38672  
 A:Molecule type: DNA  
 A:Residues: 1-29 <HA>  
 A:Cross-references: UNIPARC:UPI000016A882; GB:M38425; NID:g181977; PIDN:AAA63171.1; PID:  
 R:Xu, Y.; Iehli, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.; Mer  
 Nature 309, 806-810, 1984  
 A>Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs  
 A:Reference number: A00642; MUID:84245835; PMID:6330563  
 A:Accession: A00642  
 A:Molecule type: mRNA  
 A:Residues: 150-187, 'KSVIQA', 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-321  
 A:Residues: 713-964 <LIN>  
 A:Cross-references: UNIPARC:UPI00001725BD  
 A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF recep  
 R:Lin, C.R.; Chen, W.S.; Krulliger, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verma, I.  
 Science 224, 843-848, 1984  
 A>Title: Expression cloning of human EGF receptor complementary DNA: gene amplification  
 A:Reference number: A43615; MUID:84196372; PMID:6326261  
 A:Accession: A43615  
 A:Molecule type: mRNA  
 A:Residues: 713-964 <LIN>  
 A:Cross-references: UNIPARC:UPI00001725BF  
 A:Experimental source: epidermal carcinoma cell line A431  
 R:Stamen, F.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.  
 Biochem. Biophys. Res. Commun. 124, 125-132, 1984  
 A:Reference number: A23062; MUID:85046483; PMID:6093780  
 A:Accession: A23062  
 A:Molecule type: mRNA  
 A:Residues: 1028-1210 <SIM>  
 A:Cross-references: UNIPARC:UPI00001725BF  
 R:Weber, W.; Gull, G.N.; Speiser, J.  
 Science 224, 294-297, 1984  
 A:Reference number: A05281; MUID:84172183; PMID:6324343  
 A:Accession: A05281  
 A:Molecule type: protein  
 A:Residues: 25-30, 'S', 32-51, 454-467 <WEB>  
 A:Cross-references: UNIPARC:UPI00001725C0; UNIPARC:UPI00001725C1  
 R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Scarso, J.V.  
 J. Biol. Chem. 260, 5205-5208, 1985  
 A>Title: Identification of residues in the nucleotide binding site of the epidermal grow  
 A:Reference number: A60143; MUID:85182650; PMID:2985580  
 A:Accession: A60143  
 A:Molecule type: protein  
 A:Residues: 740-744, 'X', 746-747 <RUS>

A/Cross-references: UNIPARC:UPI00001725C2  
 R:McCzokewski, B.; Mosig, G.; Cohen, S.  
 Nature 309, 270-273, 1984  
 A>Title: ATP-stimulated interaction between epidermal growth factor receptor and superec  
 A/Reference number: A38023; MUID:8419154; PMID:6325948  
 A/Contents: annotation; receptor activity  
 A/Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA  
 R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.  
 Cell 59, 33-43, 1989  
 A>Title: Functional independence of the epidermal growth factor receptor from a domain t  
 A/Reference number: A33331; MUID:90003233; PMID:2790960  
 A/Contents: annotation; internalization signal  
 C/Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor c  
 A/Gene: GDB:EGFR  
 A/Cross-references: GDB:120610; OMIM:131550  
 A/Map position: 7p12.3-7p12.1  
 C/Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology  
 C/Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
 F:1-24/Domain: signal sequence #status predicted <Sig>  
 F:25-1210/Product: EGF receptor #status predicted <NMT>  
 F:25-645/Domain: extracellular #status predicted <EXT>  
 F:75-300/Domain: EGF receptor extracellular domain repeat <EE1>  
 F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>  
 F:646-668/Domain: transmembrane #status predicted <TM>  
 F:669-1210/Domain: intracellular #status predicted <INT>  
 F:710-978/Domain: protein kinase homology <KIN>  
 F:718-726/Region: protein kinase ATP-binding motif  
 F:999-1046/Region: coated-pit mediated internalization signal  
 F:1047-1210/Region: inhibitory  
 F:128-175,352,413,444,528/Binding site: carbohydrate (Asn) (covalent) #status predic  
 F:745/Active site: Lys #status experimental

Query Match 46.5%; Score 3169; DB 1; Length 1210;  
 Best Local Similarity 49.8%; Pred. No. 3,8e-125;  
 Matches 632; Conservative 177; Mismatches 345; Indels 116; Gaps 22;

```

11 LLLALLPREGA--STQVCGTDMKRLRPAPEPTHLDMLRHLYOGGQVGNELTYPTN 68
14 LLAALCPASRALEBKVKVCGSTSNKLTQLCTFEDHFLSLDRMNCEVVLGNLEITYVQRN 73
69 ASLSFLDIOIOVQGVYLIAHQVROVPLQRLRIVGTQLFEDNYALAVLNDGPIANNTTP 128
74 YDLSFLKTIQEVAGVYLIALNTVERILENLIQIGNMVYENSVALAVLND----- 126
129 VTGASPGGIRELPOLASLTLEILKGVLIQRNPOLCTQDTTILMKDTRHKNQALTLIDTKR 188
127 ---ANKTGLKELPMRNLQRIILGAVRFNSNNPALCVNESIQMRDIVSSDFLSNMSDFQNH 183
189 SRAICPCSPMCKSGSCWSSSEDCSLRTTVCAAGCA-RCKGRLPTDCCHEGCAAGCTGP 247
184 LGSCKKCDSCPNCGSGAGSENCQCLYKICAOQCSGRCKRSPDCCANCAAGCTGP 243
248 KHSDDLACIAHPNHSICELHCPALVTYNTDTESMPNPEGRYTFGASCVTACPYNYLSTD 307
244 RSDSLVCRKPFDEATCKOTCEPMLYNTTYQMDVNPBGKXSFATCTCKKPRRYVYTD 303
308 VGSCTLVCPRLHQVTAEDGTORCEKSPCARVCYGLAMEHLREYRAVTSANIDFPAC 367
304 HSSCRACGADSYEM-EEBDGVAKCKKCEGPRCKVCGIGIGFQKSLISINATNIKGFKNC 362
368 KIFPSLAFPSFQDDPASPANLPLOPBOLOVPELTLEETGLVLYISAWDDSLPDISVPO 427
363 TSIISDHLHLPVAFPGDSTFTHTPDPOLPDLTKVKEITGLVLLIQAMPEKNTDHLAFEN 422
428 LQVIRGRILHNGAVSYLTQGLISWLGSLRELGSGLALIHNNHTLCFVHTVPMDQLFR 487
423 LEIIRGRTQHQHOFSLAVSLNITSLGRLSLKEISGDVYIIGNNKLCANTINNKKLF 482
488 NPHQALLTANRPEDECVGEGALCHQLCARHGHWGPFPQCVNCSQFLRGQECVEBCRYL 547
483 TSGQTKTISNRGENSCKATQGVCHALCSPEGCWGPBPDPVSCNRVSGREGRCVCKKLL 542

```

```

548 QGLPREYVNAHRLCPHEPCOPONGSVTCFGEADQVACAHYKDPFPCVACPSGVKPD 607
543 EGEPRPEFENSECICCHEPCLPQAMNITCTGGRGPNICQACAYIDHPHCVTCKPAGVNGE 602
608 LSYPIKRPPEBEGACQCPINCHTSQVDDDKGCPAORASPLTISIISAVG---ILLY 664
603 NNTL-VMKYADAGHCHCHPRCTYGCQFGLGEGPTNGPKP-P-SITGMGALLLLLV 659
665 VLVGAVFGILLKRRQKLRKTYMRLLQETELVEPLTPSGAMPNOAORILKETELRVK 724
660 VALGIG---LPMRRRHIVRKTLRLRLDLRRLVEPLTSGAPNOLARILKETEPKIK 716
725 VLSGAFGTVYKGIWIPGSENVKIPLAIKVLRENTSPRANKIIDEAYVMAGVSPVYSR 784
717 VLGSAGFTVYKGIWIPGSENVKIPLAIKVLRENTSPRANKIIDEAYVMAGVSPVYSR 776
785 LIGILTSTVOLVQLMFYGCLLDHVRNKGRLSGODLNNCMQJAKMSYLEDVRLVHR 844
777 LLGILTSTVOLVQLMFYGCLLDHVRNKGRLSGODLNNCMQJAKMSYLEDVRLVHR 836
845 DLAAENVLYKSPNHYKINDPGLARLLDDITPEYHADGGKVPKMMALSIARRPTHOSD 904
837 DLAAENVLYKTPQHVKITDPGLAKLLGABEKRYHAAGKVPKMMALSIARRPTHOSD 896
905 VMSYGVTVWELMTFGAKPYDGIPLAREIPDLLEKGRRLPQPICTIDVYIMVKKCMWIDSE 964
897 VMSYGVTVWELMTFGSKPYDGIPLAREIPDLLEKGRRLPQPICTIDVYIMVKKCMWIDAD 956
965 CRPRFRELIVSESRMARDPQRFVVIQ-NEBQAPASPLDSTFYRSLIEDDKGDLVDABEY 1023
957 SRPKFREILIEFSKNAARDPQRFVVIQ-NEBQAPASPLDSTFYRSLIEDDKGDLVDABEY 1016
1024 LVPOGFCPRDPAFAGSMVHNHRSSTSGGGDLTGLSEPSSEARSPAPASEGAGS 1083
1017 LIPQGGF-----SSGSTRTPLSLSLSTS 1042
1084 DVFDGDLGMAKGLQSLPTDPSPLQRYJEDPTVPLPSET--DGYVAPALTCSPQPEYVN 1141
1043 N-INSTVACIDRNGQSPICRIBDSFLQRYSDPTALNEDSIDTFLL-----PVPEYIN 1094
1142 QPVRPQPPSPREGPLPAPRAGATL-----ERPTLSPGKGVKGVKVPAPAGAVENBEY 1196
1095 Q-----SVP-KRPAQSVQNPVYHNQPLNAPSPSDDPHYD--PHSTAGNBEY 1138
1197 L-TPQGAAPQHPRAPSPAFDNLVYWDQ-----DP-----PERGAPSTPKG 1239
1139 LNTVQ-----PTCVNSTFSDPAHMAQSGHQLSLDNPYQDDFPPEAKKPGVIRKG 1189
1240 TPTAENPEYL 1249
1190 S-TAENAEYL 1198

```

RESULT 5  
 A3183  
 epidermal growth factor receptor precursor mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 05-Oct-2004  
 C/Accession: A53183; A43818; S24942; A28941; S45325; I49643  
 R:Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Barp, H.S.; Jenkins, N.A.;  
 Genes Dev. 8, 399-413, 1994  
 A>Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor t  
 A/Reference number: A53183; MUID:94170986; PMID:8125255  
 A/Accession: A53183  
 A/Molecule type: mRNA  
 A/Residues: 1-1210 <LDB>  
 A/Cross-references: UNIPROT:Q01279; UNIPARC:UPI0000175614; GB:U03425  
 R:Aviv, A.; Lax, I.; Ullrich, A.; Schlesinger, J.; Givol, D.; Morée, B.  
 Oncogene 6, 673-676, 1991  
 A>Title: Comparison of EGF receptor sequences as a guide to study the ligand binding sit  
 A/Reference number: A43818; MUID:91232866; PMID:2030916  
 A/Accession: A43818  
 A/Molecule type: mRNA

A:Residues: 1-714 <AVI>  
A:Cross-references: UNIPARC:UPI0000175615; GB:X59698  
R:Eisinger, D.P.; Serreio, G.  
submitted to the EMBL Data Library, June 1992  
A:Reference number: S24942  
A:Accession: S24942  
A:Molecule type: mRNA  
A:Residues: 969-971, 'K', 973-1115, 'D' <EIS>  
A:Cross-references: UNIPARC:UPI0000175616; EMBL:Z12608  
R:Heisermann, G.J.; Gill, G.N.  
J. Biol. Chem. 263, 13152-13158, 1988  
A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylated  
A:Reference number: A28941; MUID:88330814; PMID:3138233  
A:Accession: A28941  
A:Molecule type: protein  
A:Residues: 689-694, 'X', 696-704, 'L', 706-707, 989-992, 'XX', 995-996, 'X', 998-1000, 1002-1009,  
A:Cross-references: UNIPARC:UPI0000175617; UNIPARC:UPI0000175618; UNIPARC:UPI0000175619;  
R:Hibbs, M.L.; Dunn, A.R.; Alexander, W.S.  
submitted to the EMBL Data Library, April 1994  
A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor  
A:Reference number: S45325  
A:Accession: S45325  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-971, 'K', 973-1210 <VER>  
A:Cross-references: UNIPARC:UPI00002182B; EMBL:X78987; NID:9488830; PIDN:CA55587.1; PI  
R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Dey, S.K.  
Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993  
A:Title: Expression of the epidermal growth factor receptor gene is regulated in mouse b  
A:Reference number: 149643; MUID:93126380; PMID:7678348  
A:Accession: 149643  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 12-20, 22-132 <RES>  
A:Cross-references: UNIPARC:UPI000016C026; GB:I06864; NID:9193001; PIDN:AA53029.1; PID:  
A:Gene: EGFR  
C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology  
C:Keywords: ATP, growth factor receptor; kinase-related transforming protein; phosphoprc  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:648-670/Domain: transmembrane #status predicted <TM>  
F:712-977/Domain: protein kinase homology <KIN>  
F:720-728/Region: protein kinase ATP-binding motif  
F:680, 695/Binding site: phosphate (Thr) (covalent) #status experimental  
F:697, 1070, 1071/Binding site: phosphate (Ser) (covalent) #status experimental  
F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental  
F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental  
F:1157/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 46.1%; Score 3145; DB 2; Length 1210;  
Best Local Similarity 49.7%; Pred. No. 3, 8e-124;  
Matches 635; Conservative 169; Mismatches 353; Indels 120; Gaps 24;

QY 11 LLLALLPFGAA--STQVCTGTDMLRLPASPTLHMLRLHYQGGVQGNIELTYLPTN 68  
DB 14 LLLTALCAAGALAEKKVCGSTSNRLTOLGTFFEDHFLSLQRMYNCEVVLGNLEITYVORN 73  
QY 69 ASLSFLDIOIEQVGVLLAHNOVROVPLORLRIYVGTOLFEDNVALLAVLDNGDPLNNTTP 128  
DB 74 YDLSLKLTIOEVAGVLLALNTVERIPLENLDIINGNALYENTYVALILSN----- 124  
QY 129 VTGASPGGLRELQLSLTEILKGVLLIQRNPOLCYODTILMKDI-----FHKNQULATLTI 184  
DB 125 -YGTNRGTGRRLPMNLGELLIGAVRFSNNPILCMMDTIQWRDIQVNFMSMSMDL--- 180  
QY 185 DTNRSRACHPSGPMKSGRSGESSESDQSILRTYCAAGCA-RCGPLPLTDCHEQCAAG 243  
DB 181 -QSHPSGCPKCDPSGPNSSCWGSGGECCKLTKIICAAQCCSHRCGRSPSDCHNQCAAG 239  
QY 244 CTGPHSGTCLACTHNHSGICELHCPALVTYNTDFEASNPNRGRYTFGASCVTACPYNY 303  
DB 240 CTGPRPSDCLVQKQFODEATCKOTCPRLMLNPTTYQMDVNPBGKISFGATVCVKCPKNY 299

QY 304 LSTDVSGCTVCPRLHNOEVTAEQSTORCEKSKPCARVYGLGMEHLREVAVTSANTIOE 363  
DB 300 VVTDHSGSVCRACGPDYEV--EEDGIRKCKCDKGCPRKCNIGIGEFDTISATINATIKH 358  
QY 364 FAGCKTIFGSLAFIPESFPDGPASNTAPLOEQOVERTLEITGYLYISAMPSTPLS 423  
DB 359 FKCYTASGDIHILPVAFKDSFRTPTPLDRELIELTIVEITVETLITQAMPDMWIDLH 418  
QY 424 VFONLQVIRGRILHNGAVSLTLQGLISWLGRLRELSGSLALIHNTHLCFYVTPWD 483  
DB 419 AFENLEILIRGTRKHQGEPSLAVGLNITSLGRSLKETSDDVDYISGRNLCYANTIMWK 478  
QY 484 QLFENPHQALLHTANRPDEBCVSGELACHQICARHCWGPEPTCCVNCOSQLRQCEYEE 543  
DB 479 KLFCTPNQTKIMNNRBAKDCKAVNHNCPILCSSESGCGPEPRPCVSCQVNSRRGCEYK 538  
QY 544 CRVLQGLPREVYNAHCLPCHPECOFQNGSVTTCGPRAEDOCVACAHYDPPFCVACRCSG 603  
DB 539 CNLIEGEERFVENSECTIQCHPECLPQAMNITCGRGPDNCTIQCAHYIDGPHCVKTCFAG 598  
QY 604 VKPDLSTYPIKFPDEEGACQPCPINCTHSCVDLDDKGCSPAEORASPLTISIISAVGILL 663  
DB 599 IMGEMNTL-VKTKADANNVCHLCHANCTYGCAGRGLOCEVWPSGPKIPSLATIGVGLL 657  
QY 664 VVLGVVFGI-LIRROQKIRKTYMRRLLOETLEVBELPTSGAMPNOQMILKTEBLRK 722  
DB 658 FIVV-VALGIGLFFMRRRHIVAKTRLRLLORELEVEPLTPSGEAPNQAHILKTEPEKK 716  
QY 723 VKVLGSGAFGVYVGIWIMPGSENVKIPALIVLEENSPKANKILDEAYVMAGVGSYV 782  
DB 717 IKVLGSGAFGVYVGLMIPSEKVKIPALIKELBAVSPKANKILDEAYVMASVDNHNH 776  
QY 783 SRLIGICTSTVQQLVTOQLMPYGCULLDHYRBNRGRISODLLMNCQIAGKSYLEDVLY 842  
DB 777 CRLLGICTSTVQQLITQLMYGCULLDHYRBNRGRISODLLMNCQIAGKSYLEDVLY 836  
QY 843 HRDLAARNVTVKSPNNHYKITDFGLARLLIDETIYHADGCVPIKMMALBSILRRPFTHQ 902  
DB 837 HRDLAARNVTVKTPQHYKITDFGLAKLLGAEKEYHAEKGKVPKMMALBSILRRPFTHQ 896  
QY 903 SDVWSYGVTVWELMTFGKAPVDGI PAEIRPDLKGBELPQPICTIDVYIMVCKMID 962  
DB 897 SDVWSYGVTVWELMTFGSKPVDGI PASDISLEIKGBELPQPICTIDVYIMVCKMID 956  
QY 963 SECPREFELVSEFSRMAPDQRFQFVLIQ-NBDLGPASLSTFYRSLLEDMDMDVDAE 1021  
DB 957 ADSRPFRELILERSQWABDQRYLVIGQDGRMLPSTDSNFYRALMDESDMEDVDAD 1016  
QY 1022 EYLVPOQGFPCPDPAFGAGKVVHHRHSSSTRSGGDLTLGLPSEBEAPRSPAPSGA 1081  
DB 1017 EYLVPOQGF-----NSPST-----SRTPLLSLSA 1042  
QY 1082 GSDVFPDDDLGMAKGLISLPTHDSPLOFSEDPVLPSET--DGVAVALTSCPOPEY 1139  
DB 1043 TSN-----NSTVACINRNGSCRVKEDAFQRYSSPFTGAVVEDNIDDAVL-----PVPEY 1092  
QY 1140 VNQPDVAPQPPSPREBGLPAARPAATL-----BRPTLSFGKGVADVFAFGAVENP 1194  
DB 1093 VNQ-----SVF-KRPAQSVQNPVYHNNQRLHRAGRDLHQN--PHSAVGNP 1136  
QY 1195 EYL-TPQGAAPQHPPPAPSPAFDNLVYWDQ-----DP-----BERGAPSTF 1237  
DB 1137 EYLVTAQ-----PTCLSSGFSNPSALMTQKSHQMSLDNPYQODFPFKETKFPNGIF 1187  
QY 1238 KGTPTAENPEYLGIVP 1254  
DB 1188 KG-PTAENAEYLRVAPP 1203

RESULT 6  
tychiv  
epidermal growth factor receptor precursor - chicken  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbb

C.Species: Gallus gallus (chicken)  
 C.Date: 28-Feb-1986 #accession\_revision 05-May-1995 #text\_change 05-Oct-2004  
 C.Accession: A27720, A00643  
 R.Lax, I.; Johnson, A.; Hawk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Vennart  
 Mol. Cell. Biol. 8, 1970-1978, 1988  
 A.Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in mou  
 A.Reference number: A27720, M01D:88261272, PMID:3260329  
 A.Accession: A27720  
 A.Molecule type: mRNA  
 A.Residues: 1-1223 <LAX>  
 A.Cross-references: UNIPROT:P00534; UNIPARC:UPI00001725C3; GB:M20386  
 R.Nielsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rotman, F.M.; Crittenden, L.B.; Raines, M.  
 Cell 41, 719-726, 1985  
 A.Title: c-erbB activation in AIV-induced erythroblastosis: novel RNA processing and pro  
 A.Reference number: A00643; M01D:85228222; PMID:2988784  
 A.Accession: A00643  
 A.Molecule type: mRNA  
 A.Residues: 585-1223 <NIL>  
 A.Cross-references: UNIPARC:UPI00001725C4; GB:M10066  
 C.Genetics:  
 A.Gene: erbB  
 C.Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology  
 C.Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor  
 specific protein kinase  
 F.1-30/Domain: signal sequence #status predicted <Sig>  
 F.31-1223/Product: epidermal growth factor receptor #status predicted <Mat>  
 F.31-654/Domain: extracellular #status predicted <EXT>  
 F.81-107/Domain: EGF receptor extracellular domain repeat <EB1>  
 F.397-610/Domain: EGF receptor extracellular domain repeat <EB2>  
 F.655-677/Domain: transmembrane #status predicted <TM>  
 F.678-1223/Domain: intracellular #status predicted <INT>  
 F.719-984/Domain: protein kinase homology <KIN>  
 F.727-735/Region: protein kinase ATP-binding motif  
 F.136,202,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #  
 F.192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted  
 F.687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F.754/Active site: Lys #status predicted  
 F.1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match 45.8%; Score 3122.5; DB 1; Length 1223;

Best Local Similarity 48.7%; Pred. No. 3.3e-123;

Matches 632; Conservative 105; Mismatches 345; Indels 145; Gaps 25;

8 RKGGLLALPRGA-----STOVCTGDMKRLRPSPTHLDMLKQCGOVGNLE 61  
 13 KCAAVLVLLLVALLGALCSAVEBEKKVCGSTNNKLTQLGHVEDHTSLQRMVNCCEVLSNLE 72  
 62 LTYLPTNAGLSFADIQEYGVYLAHNOVRQVPLQRLRIVRGTOLEFNALVALVNDGD 121  
 73 IYVEHNRDLTLTKITQEVAGYVLLANVVDVPLENLOIRGNVLYDVSFALAVSNYH 132  
 122 PLNNTPTVTGASPGGLRELQSLTEILKGVLIQRPOLCYODTLTKDIFHKNNQAL 181  
 133 -NNKQ-----GLRELPMKRLSEILNGVXISNNPKLCMMMDVTLAMDIDTSRK-PL 182  
 182 TLID-TNRBACHPCSPMKGSRKMGSESDQSLRTYCAAGCA-RCKGRLPTDCHQ 239  
 183 TYLDPAFASNLSSCPKCPNCTEDHCWAGBONCOTLKVI CAQCCSGCRGKVPSSDCN 242  
 240 CAAGCTGPHSDCLACLNHNSGICELHPALVYTTDFEESPMREGRYTFGASCVTAC 299  
 243 CAAGCTGPRESDCLACRKRDATACTDTPRLVLYNPTTYQNDVNEBGRKYSFGATVRC 302  
 300 PNYLSTDVGSCTLVCPRLNQEVTADGTQRCCKSPCARVCGYGLMEHLREVAVTSA 359  
 303 PHNYVTVDHGSCYRSNCTDTVEY-ENNGVAKCKCKGCLSKVCGNGIGIGELKILSINAT 361  
 360 NIQEPAGCKKIRGSLAFPESPRGDPASTAPYQRPOLQVETLEITGVLYSAMPDL 419  
 362 NIDSFNCKTKINDVAILPAFAGDFTKTLPLDPKKLDVFRVKSISGLLIQAMPDVA 421  
 420 PDLSPQNTQVTRGRILHNGAVSLTQGLGISWLGRLSRLSGLALIHNTHLCFVAT 479

422 TDLYAFENLEIRGRTHQGOYSLAVNLKISLGLSLKELISDGDIAIMKNKLCYADT 481  
 480 VPMDOLEFRPHQALHTANRPEDECVSGSLACHOLCARGHCWGPPTCVNCSQFLRGE 539  
 482 KMRSLFATQSOKTKTIONRKNDDCADHVCDDPLCSDVCGKGPFGFCSCFRFSRKE 541  
 540 CYEECGVLOGLPREYVNAHCLPCHPECCPONG---SVTCFGRPADQVCAAHYKDPFC 596  
 542 CYKQCNIIQGEPRFRERDSKCLPCHSECLVQSTAVNTTCSGPGPDHCKKCAHFDGPHC 601  
 597 VARCSGVKPDLSYMPWKPEDEBGAQCPPINCTHSQVLDLDDKCPAEQASPLTSTIS 656  
 602 VACAPAGVLGENDTL-VMKYADANAVCOLCHENCTRCGCPGLEGP---NSSKTPSITA 657  
 657 AVV-GILVVYLVGVFGILIKRROCKIRKYMTRBLLOEYELVPLTPSGAMPNOQMRL 715  
 658 GAVGGLCLVAVGLGGLYLRRL-HVRRKTLRLLOEYELVPLTPSGEAPNOAHKRL 716  
 716 KETELRKVVLGSGAFGVYKGIWIPDGENVKIPVALKYLRNTSPKANKELIDEAYVNA 775  
 717 KETEPKRVKVLGSGAFGVYKGLMIPGEKVKIPVALKELRATSPKANKELIDEAYVNA 776  
 776 GVSGPVYSLIGICTSTVQVLTQMPYCGLLDHVRENRGLSGODLWMCQIAKMSY 835  
 777 SYDNPHVCRLGICTSTVQVLTQMPYCGLLDIYREHNDIGSOYLMMCVQIAKMSY 836  
 836 LBDVLYPHDIAARVLYKSPHNVKITPDGLARLIDIDTEYHAGSKPIKMMLESL 895  
 837 LERRLYPHDIAARVLYKSPHNVKITPDGLARLIDIDTEYHAGSKPIKMMLESL 896  
 896 RRRFTQSDVMSYGVYVWELMTFGAKPYDGI PARERIDLEKGERLPORPCTIDVYMT 955  
 897 HRIYTHQSDVMSYGVYVWELMTFGSKPYDGI PARERIDLEKGERLPORPCTIDVYMT 956  
 956 VKCMNIDSCRPRFELVSEFSRMAPDQRFVYIQ-NEDLGPASPLDSTFYSLLEDMM 1014  
 957 VKCMNIDASRPKRELIAFSGMKARDPRVLYIQDEBEMHLPSPDTSKFYTYLMEEDM 1016  
 1015 GDLNDAERTLVPOGFPCDDPAPGAGGVYHHHRSSSTSGGDLTLGLPBBEAPRSP 1074  
 1017 EDIVDAEYLVHVGPF-----NSPST-----SKRP 1042  
 1075 L-----APSEGASDVFDGDLGMAKGLQSLPTHDPSPLQYSEDPVPLPSET--DGY 1127  
 1043 LLSLSATSNNSATNCID-----RNGCHPVAEDSVQYSSDPTNPLSESIDDF 1094  
 1128 VAPLTCSPOPEYVNOQDVARPOPSPREGELPARAPAGATLERPKTSLSPKNGVXDF-- 1185  
 1095 L-----PAPEYVNO--LMPKKPS-----TAMVQNOIYNNISLT 1125  
 1186 -----AEGAVENREYLTPOGDAAPQHPPPAPAPDNLVYMQ----- 1225  
 1126 AISKLPMSRYSNASTADNPEYL-----NTNQSPLAKTYESSSPYMIQSGNHQIN 1177  
 1226 -DPE-----RGAPSTFKGPTPAENRYGLDVP 1254  
 1178 LNPDIQDDFLNRTKFNGLVPAENPEYLRVAAP 1214

RESULT 7  
 A47253  
 C.Species: Homo sapiens (man)  
 C.Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Oct-2004  
 C.Accession: A47253  
 R.Plowman, G.D.; Clouscou, J.M.; Whitney, G.S.; Green, J.M.; Carlton, G.W.; Foy, L.; Ne  
 Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993  
 A.Title: Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal  
 A.Reference number: A47253; M01D:93189574; PMID:8383326  
 A.Accession: A47253  
 A.Molecule type: preliminary; not compared with conceptual translation  
 A.Status: preliminary; not compared with conceptual translation  
 A.Reference number: A47253; M01D:93189574; PMID:8383326  
 A.Molecule type: nucleic acid  
 A.Residues: 1-1308 <PLOW>



```

60 LETVLPNTASLSPLODIOEVGVYLIAHNOVQVPLQRLIVRGTOLEFEDNYALVDN 119
63 LEITTOENODISFLOSIQIEGVGVLIANNEVSTIPLVNLRILRGONLIEGNFTLLIMSN 122
120 GDPINNTTPVGTASPGELRELOLRSLTEILKSGVLIQBNPOLCYODTILMKDIFHKNNOL 179
123 YQK-NPSSP--DVYQGLKQLQLSNLTIELSGGVKSHNPLLCANETIMWMDIVDKTSP 179
180 ALTLIDTRNSRACHCSPKCKSRCKGSESDCOSLTRVCGAGC-ARKGGLPDDCCHE 238
180 TNNLLPHAFEROCOCODHCNVGSCWAPRGHCKFTKLLCAEQCNRRRGKPDCCNE 239
239 OCAAGCTGPKISDCLACHFNHSGICELHCPALVTNTDTPESMPREGYFGASCVA 298
240 HCAAGCTGPRATDCLACRFNDGTCCKTCCPPKTYDIYSHQVNDPNKTYFGACVKE 299
299 CPYNYLSTDVGSCTLVCPLANOEVTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTS 358
300 CPSNVVYTE-GACVRSACAGMLEVD-ENKRSCKPCDGVCPKVCDDIGIGLSNTLAVNS 357
359 ANIQFPAGCKKIFGSLATLPESFDGDDPASNTAPLOEQLQVETIABETIGYLYISAMPDS 418
358 TNIRSFNCTKINGDIILNRNSPEGDPHYKIGTMDPEHLMNLTYSKELTGYLVIMWMPEN 417
419 LPDLSVPONLQVIRGRILHNGAYS-LTLOGIGISWLGRLSRLSGLALIHNNHLCPV 477
418 MTSLSVFPOLLETIRRTTFSRGFSFVVQVNRHLQWGLSLKESVAGNVILKNTLQLKRA 477
478 HTVPMDQFLRNPHQALLHTANRPEDECEVGEGLACHQLCARGHCMWPGFPTQCVNCSQPLRG 537
478 NTIMNRRLPRSEDOIEVDART-----ENQTCNNESEDDCM-EGPTMCVSCILHVRG 529
538 QCEVECEVLOGLPREYVNAHCLPCHPECCOPQNSVTCFGEADQCAACAHYKDPFPCV 597
530 GRCVASCNMLQOEPRBAQVGRCVQCHOECLVQTSLTCYGPANCSKSAFQCGPQCI 589
598 ARCPGVPRLSYMVIKWPDEEGACOPCPICTHSCVDLDDKGPAGQASPLSIISA 657
590 PRCPHGLIGDGT-LIMKADMMGOCQPHQNCQCCSGPGLSGRGD-IYSHSLAVGL 647
658 VVGLILVVVLGVFGILIKRQOKIRKYTMRLQETELVEPLTPSGAMPNOQMRILKE 717
648 VSGILITVIALLVILVLRRIK-RKQIRCLLDEKELVEPLTSGAPNQAFLRIKE 706
718 TELRKVKVLGSGAFGTVYKGIWIPGENVKIPVAIKVRENTSPANKELIDEAYVMAV 777
707 TEFKDRVLGSGAFGTVYKGLMNPDEGEMRIPVAIKVRENTSPANKELIDEAYVMAV 766
778 GSPVSRLLGICLTSTVOLVQMLPYGCLLDHVRNBRGLSGODLLNMCQOIAKMSYLE 837
767 DHPHCRLLGICLTSAVOLVQMLPYGCLLDHVRNBRGLSGODLLNMCQOIAKMSYLE 826
838 DVRLVHRDLAANVAVKSPNHVKTIDPGLARLLIDETEHADGGKVPKMMALSEIIR 897
827 ERHLVHRDLAANVAVKSPNHVKTIDPGLARLLIDETEHADGGKVPKMMALSEIIR 886
898 RPTHOSDWSVGVYTWELMTFGAKPYDGIIPAREIPDLLEKGRLLPQPPICITDVYMIWK 957
887 TYTHOSDWSVGVYTWELMTFGSKPYDGIIPAREIPDLLEKGRLLPQPPICITDVYMIWK 946
958 CMMIDSEBRPRELVESESRMARDDORFVVIONDLPASPLDSTFYRSLIEDDDMDL 1017
947 CMMIDSEBRPRELVESESRMARDDORFVVIONDLPASPLDSTFYRSLIEDDDMDL 1001
1018 VDABEYLVPQOGFCPPDPAPGAGVNHHRSSSTRSGGDLTLGLLEPSSEBRPSPLAP 1077
1002 VDADHYLVPYKRI-----NRQSS-----EPCIP 1024
1078 SEGAGSDVFPDGLGMAKAGLOSLPTHDBSPLOARYSEDPV-PLPSEPDGYVAPLTCSPO 1136
1025 PTGH-----PVRENSITLRNISDPTQNALREKDLGH----- 1055

```

```

QY 1137 PEYVQPPVRPQP-----PSPRE-----GELP-AARPAAGATLRRPKTLSPKNGVVD 1183
DB 1056 -EYVQPPGSESSRSRLDIYNPYEDLTDGKGVSLSSQCAETNFRPELTANNQSL----- 1111
QY 1184 VFAFGAVENPERYLTTPGGAAAPQHPPPAFSPAFNLYYWDODPEPRGAPSTFFKTPPA 1243
DB 1112 PLVSSGSMDDPDY---QAG-----YQAAAF-----LPQTALTGNGMFLPAA 1149
QY 1244 ENPEYLG 1250
DB 1150 ENPEYLG 1156

RESULT 9
A36223
kinase-related transforming protein (erbB3) (EC 2.7.1.-) precursor - human
C:Species: Homo sapiens (man)
C:Date: 04-Oct-1991 #sequence_revision 13-Jan-1993 #text_change 31-Dec-2004
C:Accession: A36223; 159164
R:Kraus, M.H.; Issing, W.; Miki, T.; Popescu, N.C.; Aaronson, S.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 9193-9197, 1989
A:Title: Isolation and characterization of ERBB3, a third member of the ERBB/epidermal g
A:Reference number: A36223; MUID:90083234; PMID:2687875
A:Accession: A36223
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1342 <KRA>
A:Cross-References: UNIPROT:P21860; UNIPARC:UPI000017A3AE; GB:M29366
R:Plowman, G.D.; Whitley, G.S.; Neuhauer, M.G.; Green, J.M.; McDonald, V.L.; Todaro, G.J
Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990
A:Title: Molecular cloning and expression of another epidermal growth factor receptor-re
A:Reference number: 159164; MUID:90311312; PMID:2164210
A:Accession: 159164
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-559, 'G', 561-957, 'P', 959-1063, 'G', 1065-1342 <RES>
A:Cross-References: UNIPARC:UPI000050P2D; GB:M34309; NID:9183990; PID:AAA5979.1; PID
A:Genetics:
A:Gene: GDB:ERBB3; HER3
A:Cross-References: GDB:119880; OMIM:190151
A:Map position: 12q13-12q13
A:Keywords: ATP; phosphotransferase
F:707-972/Domain: protein kinase homology <Kin>
F:715-723/Region: protein kinase ATP-binding motif

Query Match 35.7%; Score 2432.5; DB 2; Length 1342;
Best Local Similarity 40.7%; Pred. No. 2,2e-94;
Matches 533; Conservative 191; Mismatches 458; Indels 129; Gaps 32;

QY 10 GILLALALPPGA--STVCTGDMKLRLPASPETHLDMRLHYOGCOVVOGRLVLTLP 67
DB 11 GLPSLARGSEYNSQAVCPGLTNGLSVTDANEOYQLTKYLKCEVVMGNLEIVLGH 70
QY 68 NASISFLDIOEVGVYLIAHNOVQVPLQRLIVRGTOLEFEDNYALVDNGBLANTT 127
DB 71 NADISFLQWIREVSYLVANNEFSTLPLPMLRVVRGQYVDKRAIFVM-----LNTVT 125
QY 128 PYTGASPGELRELOLRSLTEILKSGVLIQBNPOLCYODTILMKDIFHKNNOLALITDN 187
DB 126 ----NSSHALQRLTQLTIELSGGVYIEKNDKCMQDITIMRDIRDRD--ABIYVD 178
QY 188 RSRACHPSPKCKSRCKGSESDCOSLTRVCGAGC-ARKGGLPPTCCHEQCAAGCTG 246
DB 179 NGRSCPCHAYCKG-RCKGPGSEDDQTLTKITICAPQCNHCGFPNPNCCCHDECAAGGSG 237
QY 247 PKHSDCLACHFNHSGICELHCPALVTNTDTPESMPREGYFGASCVTAAPYNYLST 306
DB 238 PQDTCFACRHPHNSGACVPRCPQPLVYNNKLTFLQLEBPHTKYGVGVASCPHNFV-V 296
QY 307 DVGSCTLVCPLANOEVTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSANIQEPAG 366
DB 297 DQTSVCRACPPDKMEVD-KNGLKMEPCGGGLCPKACBGTGGG--SRFQTVSSNIDGFPV 353

```



367 CKKIFGSLAFLPESPDGDPASNTAPLOEOLQVETLEETITGYLYISAMPDLSPLDSVFO 426  
DB CTKILGNLDFLTGLGNPWHKIPALDPEKLVNFTVEITGYLYNIQSMPPHMFVSFVS 413  
427 NLQVIRGRLIHNGAVS-LTLQGLGISWLGRLSRLSGSLALIHNTLCPHTVYPMQOL 485  
414 NLTTGGSLYNGRGSLLIMKLVNTSLGFSRLKESISAGRIYISNRDLCYHHSLSMKV 473  
486 FRNPHQALLHTA-NRPEDECVGEGLAGHOLCARHGCMWPPTQCVNCSOFLRGQCEVEC 544  
474 LRGPFEERLDIKNRPDRDCAVAGKVCDBLSSGGCMGPPGQCLSCANYSRGVCVTHC 533  
545 RVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVACRPSGV 604  
534 NFLNGEPREFAHBAECFSCHEPCQMEGATCNNGSGSDTCAQCAHFRDGPCHVSSCPHGV 593  
605 KPDLSTMYIKRPDEGACQPCPTNCTHSQVLDKCGPAERAA---SPLTISAVVG 660  
594 LG-AKGPITYKYPDVQNECRPCHECTGCKGPELDQCLGQTLVLIGKTHLTMALTVIAG 651  
661 ILVVVLGVFGILIKRRQOKIR-KYTMRRLLQETELVEPLTPSGAMPNOAMRIKETE 719  
652 --LVIFPMMLGSTFLYMGGRIONGBARRIYERGESIEPLDPS-EKANKVLARIKETE 708  
720 LRKVVILGSGAGFTVYKGIWIPDGENVKIPVAIKVLRNTSPKANKELIDEAYVNAVGVGS 779  
780 PYSVSLILGICLTSTVQVLTQMLPYGCLLDHRENGRLSGODLLMWCQOIKMGSYLSDV 839  
769 AHIYVLGLCPRESSIQLVLYQLPGLSLDHNQHGALGAPOLLNMGVQIAGMYLSEH 828  
840 RLVRHDLAARNVLVSPNHVKITDFGLARLIDIDEYEADGCKVPIKMALESILRRRF 899  
829 GMYHNNLAARNVLKSPSQGVADFGVADILRPDKQOLLYSEAKPIRIMMLESIHPEKY 888  
900 THQSDVMSYGVTVWELMTFGAKPYDGIAPAREIPDLLEKGERLPPPTCTIDVYMIYKCM 959  
889 THQSDVMSYGVTVWELMTFGAEPYAGRLAEVDPDLLEKGERLAPQICTIDVYMMVYKCM 948  
960 MIDSCRPRFELVSEFRMADPPRFVYIOMEDIGRA---SPLDSTYRSLLEDDMGD 1016  
949 MIDENIRTEKELANEFTRMADPPRYLVIKES-GPGIAPGEPHGLTNKKLEVELELP 1007  
1017 LVDAEYVLVPOQGFPCPDPAFGAGMVHHRSSSTREGGGDLTLGLEP-SEEAERAPSL 1075  
1008 ELUDLDLDEABED-----NLATTTLSALSLPVGTLNRPRESSQLL 1048  
1076 APSEGASGVDPDGDLGMAKGLQSLPTHD-PSPLQRYSEDPVLP-----SETDGYV 1128  
1049 SPSGGY-MPMNQNLGESCQESAVGSSBRCPVSLH-----PMPRGCLASSESGHV 1101  
1129 A-----PLTCSPOPE-----YNNQDVNRPQPSREGE-----L 1157  
1102 TGSEAELOEKVSMCRSRSRSPRPRGSAVHSQHSLTLPTPLSPGLEBEDVNGYVM 1161  
1158 PAARAPAGATLERPKTLP-GKNGVV-----KDVAFGAAVENPEYLPPOGGAAPQAPP 1210  
1162 PDTHTKTPSSREGTLLSSVGLSVGTEBEDD-----EYVYNNRRRRHSP-PHP 1212  
QY 1211 PAFSAPFNLVYWD-----QDPERGAPSTFTKGTPTAENPEYL 1249  
DB 1213 RPSLEELGYEYMDVGDLSLSTGSCPLHPVIMPAGTTPDEDEYEM 1263

RESULT 10  
JC4387  
epidermal growth factor receptor homolog precursor - rat  
N:Alternate names: ErbB3 protein; HER3 protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 17-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 05-Oct-2004  
C:Accession: JC4387  
R:Heilayer, N.J.; Kim, H.H.; Greaves, C.H.; Sierke, S.L.; Koland, J.G.

Gene 165, 279-284, 1995  
A:Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protein.  
A:Reference number: JC4387; MUID:96096535; PMID:8522190  
A:Accession: JC4387  
A:Molecule type: mRNA  
A:Residues: 1-1339 <HEL>  
A:Cross-references: UNIPARC:UPI000017A3DA; GB:U29339; NID:9915389; PID:9915390  
A:Experimental source: liver  
A>Note: The authors translated the codon AAC for residue 369 as Thr and GTT for residue  
C:Comment: This protein is a functional heregulin receptor that transduces signals to th  
C:Genetics:  
A:Gene: ErbB3  
C:Keywords: ATP, growth factor receptor, liver, phosphoprotein, transmembrane protein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-1339/Product: epidermal growth factor homolog #status predicted <MAT>  
F:640-659/Domain: transmembrane #status predicted <TMM>  
F:705-970/Domain: protein kinase homology <KIN>  
F:713-721/Region: protein kinase ATP-binding motif  
F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr) (C  
Query Match 34.4%; Score 2347.5; DB 2; Length 1339;  
Best Local Similarity 40.8%; Pred. No. 7.8e-91;  
Matches 524; Conservative 170; Mismatches 434; Indels 155; Gaps 34;  
QY 3 LAALCRMGLLIALIPPGA---STQVCTGIDMKRLRPASEPETHLDMRLHYOGCOVQGN 59  
DB 7 LQVLC---FLSLARGSEMNSQAVCPGTINGLSVTGDANOYQTLVLYKEKEVWGN 62  
QY 60 LELVYLPNTNASTLFDIOIEQVGYVLLAHNOVRQVPLRIYVGTOLFEQNYVALAVDN 119  
DB 63 LEIVLTGNHDLSTFLQWIREYAVYLVANMFSLVPLNVLVYKSTQVYDQKFLFVN-- 120  
QY 120 GDPVNTTTPVYASPGGIREQLRLSTLEIKGVLIQRNPQCYQDTILMKDIFKNNQL 179  
DB 121 ---LNVNT---NSHMLRLQKFTQLTEILSGGYIEKNDLCHMDITDMEDIVRVR--- 170  
QY 180 ALTLIDITNRSAPCHPCSPMCGSGRCWGESSEDQSLRTVCAGGC-AACKPLPDPCCHE 238  
DB 171 GAEIVKNGNANGACPCHCVCNG-RCWGGPDPDCCILTKTICAPQCNCRCPFPNPQCHD 229  
QY 239 QCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNPRGRYTFPGASCTYA 298  
DB 230 ECAGGCGSPQDTDFACRFRPDSACVPRCEPLVYNTKLTQLERNPYTKIQYGVVCAVS 289  
QY 299 CPYNYLSTDVSGCTLVCPRLHNOEYTABDGTORCEKCSKPCARVCYGLGMEHLREYAVTS 358  
DB 290 CPNHFV-VDQFCYRACPRDMEYD-KHGLMCGSPCGGLCPKACEGTSG--SRVQTVDS 345  
QY 359 ANIOEPACKKIFGSLAFLPESPDGDPASNTAPLOEOLQVETLEETITGYLYISAMPDS 418  
DB 346 SNIGFVNTCTKILGNLDFLTGLVNDPWHKIPALDPEKLVNFTVEITGYLYNIQSMWPH 405  
QY 419 LPDLSVQNLQVIRGRLIHNGAVS-LTLQGLGISWLGRLSRLSGSLALIHNTLCPHTV 477  
DB 406 MENSVSFNSLTTTGGSLYNGRGSLLIMKLVNTSLGFSRLKESISAGRIYISNRDLCYH 465  
QY 478 HTVPMQDLFRNPHQALLHTA-NRPEDECVGEGLAGHOLCARHGCMWPPTQCVNCSOFLR 536  
DB 466 HSLMNTRLLRPSERLDIKYDRPLGECIAGKVCDBLSSGGCMGPPGQCLSCANYSR 525  
QY 537 GQECYECRVLYQGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFC 596  
DB 526 EGVGVCTHGNFQGPREFEVHQAQFSCHEPCLPMEGISTYNGSSSDACARCAHFRDGH 585  
QY 597 VAPCPGVKPLDSMPRIKRPDEGACQPCPTNCTHS--VDLDDKCGPAERAPSLRSI 654  
DB 586 VNSCPHGILG--ANGPIYKIPDAQNECRPCHECTGCKGPELDQCLGQAEVLSKPLV 643  
QY 655 ISAVGILLVVLGVFGILIKRRQOKIR-KYTMRRLLQETELVEPLTPSGAMPNOAMR 713  
DB 644 IAVYVG--LAVYIMLIGSFLYMGGRIONGBARRIYERGESIEPLDPS-EKANKVLAR 700  
QY 714 ILKETLRKVKVLGSGAGFTVYKGIWIPDGENVKIPVAIKVLRNTSPKANKELIDEAYV 773



```

Db      701  I F K E T E L R K L K V L G S G V G T V H N K G I W I E G E S I K I P V C I K T I E D K S G S G S O A V T D H L A      760
Qy      774  M A G V G S P R V S R L L G I C L T S T V Q L V T Q L M P R Y G C L L D H R E N G R L G S O L L M W C M Q I A K M      833
Db      761  V G S I D N H A I V R L L G I C P G S S I O L V T Q U L P L G S L I D H V Q H R E T L G P R O L L M W G V Q I A K M      820
Qy      834  S Y L E D V R L V H R D L A A R N V L V K S P N H V K I T D F G L A R L I D I D E T E Y H A D G K V P I K M A L E S      893
Db      821  Y L L E S H S V H R D L A R N V L K S P S O V Q V A D F G V A D L P R P D K Q L H S A K T P I K M A L E S      880
Qy      894  I L R R F T H Q S D V M S Y G V T V M E L M T F G A K P Y D G I P A R E I P D L L E K G E R L P O R P I C T I D V Y M      953
Db      881  I H F K Y T H Q S D V M S Y G V T V M E L M T F G A R P Y A G L R A E I P D L L E K G E R L A Q O I C T I D V Y M      940
Qy      954  I M V K C M I D E S C R P R F R E L V S E F S R M A R D P O R F V T I Q N E D L G P A S P L D S T Y R S I L E B D D      1013
Db      941  V M V K C M I D E N I R P F K E L A N E F T M A R D P R Y L V I K R A S - G P G T P - P A A E P S V L T T K E      997
Qy      1014  M G D L V D A E E Y L V P O G G F C P R P A R A G A G M V H N H R S S T R S G G D L T L G L E P S E E - - - - -      1068
Db      998  L - - - - - O E A L E P E L - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      1023
Qy      1069  - - - - - E A P R S P L A P S E G - - - - - A G S D V P D G L G M G A K G L S L P T H D      1105
Db      1024  L G S A L S L P T G L T R P R G S Q S L S P S S G Y M P M Q S L G R A C L D S A V L G G R E Q P R S I S L H -      1082
Qy      1106  P S P L Q R Y S E D P T V P L P S E T D G V - - - A P L - - - - - T C - - - - - S P O P E - - - - - Y N Q P D V      1145
Db      1083  P I P P G R - - - - - P A S E S S E G H V T S E A E L Q B K V S V C S R S R S R P R P R C D S A Y H S Q R H S      1135
Qy      1146  R P O P P S P R E G P - - - - - L P A R A P A G A T L E R P K T L S P - G K K G V V - - - - - K D V P A F      1187
Db      1136  L L T P T P L S P G L E E D G N G V M P D T L R G A S S R E G T L S V G L S S V L G T E B E D D - - - - -      1191.
Qy      1188  G C A V E N P E Y L T P O G G A P O P H P P      1210
Db      1192  - - - - - E E Y E Y M N R K R G S P - P R P      1209

RESULT 11
TVFVLY
protein-tyrosine kinase (EC 2.7.1.112) erbB - avian leukosis virus
N:Contains: amino end of gag protein; env protein fragment; protein-tyrosine kinase
C:Species: avian leukosis virus, ALV
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-Oct-2004
C:Accession: B00643; A00643
R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines, M
Cell 41. 719-726, 1985
A>Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pro
A:Reference number: A00643; MUID:85228222; PMID:2988784
A:Accession: B00643
A:Residues: 1-698 <NID>
A:Molecule type: mRNA
A:Cross-references: UNIPROT:P00534; UNIPARC:UPI000017113; GB:M10066; GB:M13881; NID:921
A>Note: in Genbank entry CHERBBF, release 109.0, the source is designated as Gallus gal
C:Comment: This protein is synthesized as a gag-env-erbB protein.
C:Gene: erbB
A:Gene: gag-env-erbB
C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
C:Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific p
F:1-6/Product: gag protein (fragment) #status predicted <RNG>
F:7-59/Product: env protein (fragment) #status predicted <RNV>
F:60-698/Product: protein-tyrosine kinase erbB #status predicted <ERB>
F:194-459/Domain: protein kinase homology <KIN>
F:202-210/Region: protein kinase ATP-binding motif
F:229/Active site: Lys #status predicted

Query Match      25.9%; Score 1765.5; DB 1; Length 698;
Best Local Similarity 52.2%; Pred. No. 7,76-67;
Matches 374; Conservative 80; Mismatches 137; Indels 125; Gaps 18;
Qy      578  G P E A D Q C A C A H Y K D P P C V A R C P S G V K P D L S Y M P I M K F P D E G A C O P C I N C T H S C V D L      637

```

```

Db      60  G P - - D H C M K A H F I D P H C V A C P A G V L G E N D T L - V M K Y A D A N A V C O L C H R N C T R G C G P      116
Qy      638  D D K C P R A E G R A S P L T S I T S A V - G I L L V V T L G V V F G L I K R O O K I R Y T M R L L O E T E L      696
Db      117  G L E G C P - - - N G S K T P S I A G V G G L C L V N G L G I G L T R R R - H I V R K T I R R L L Q E R L      172
Qy      697  V E P L T P S G A M P N O A M R I L K E T E L R K V L G S G A F G T Y K K I W I P D G E N V K I P V A I K Y L R      756
Db      173  V E P L T P S G A M P N O A M R I L K E T E R K K V L G S G A F G T Y K K I W I P E G K V K I P V A I K Y L R      232
Qy      757  E N T S P K A N K E I L D A Y V M A G V S P Y S R L G I C L T S T V Q L V T Q L M P R Y G C L L D H R E N G R      816
Db      233  E A T S P K A N K E I L D A Y V M A S V D N P H V C R L G I C L T S T V Q L T Q L M P R Y G C L L D Y I R E H K D N      292
Qy      817  L G S O D L M W C M Q I A K M S Y L E D V L V H R D L A A R N V L V K S P N H V K I T D F G L A R L I D I D E T E      876
Db      293  I G S O Y L L M W C V Q I A K M N Y L E E R L V H R D L A A R N V L V T P Q H V K I T D F G L A K L G A D E K E      352
Qy      877  Y H A D G K V P I K M A L E S I L R R F T H Q S D V M S Y G V T V M E L M T F G A K P Y D G I P A R E I P D L L E      936
Db      353  Y H A E G K V P I K M A L E S I L R I Y T H Q S D V M S Y G V T V M E L M T F G S K P Y D G I P A S E I S S V L E      412
Qy      937  K G E R L P O P I C T I D V Y M I M V K C M I D S E C R P R F R E L V S E F S R M A R D P O R F V I Q - N E D L G      995
Db      413  K G E R L P O P I C T I D V Y M I M V K C M I D A D S R K P R E L I A E F S K M A R D P R Y L V I Q G D E R M H      472
Qy      996  P A S P L D S T F Y R S L I E D D M G L V D A E E Y L V Q O G F P F C D P A R A G A G M V H N H R S S T P S G      1055
Db      473  L P S P L D S T F Y R T L M E E B E M E D I V D A E Y L V H Q G F - - - - - N S P S T - - -      513
Qy      1056  G G D L T L G L E P S E E A P S P L - - - - - A P E D A G S V F P D G L G M A K G L S L P T H D P S P L Q      1110
Db      514  - - - - - S R P L L S L S A T S N N A T I C I D - - - - - R N Q G H P V A E D S T V Q      550
Qy      1111  R Y S E D P T V P L P S E T - D G Y A P L T C S P O P R Y V N Q P D V R P O P S P R E G P L P A R P A G A T L E      1168
Db      551  R Y S S D P T G N F L E S I D D G F L - - - - - P A B Y V N Q - - L M P K K R S - - - - -      585
Qy      1169  R P K T L S P E K N G V A D V F - - - - - A F G A V E N P E Y L T P O G G A P O P H P P A F      1213
Db      586  - - - - - T A M Q N O I Y N N I S L T A I S K L P M D S R Y O N S H T A D N E Y L - - - - - N T N Q S P L A      633
Qy      1214  S P A F D N L Y Y M Q - - - - - D P P E - - - - - R G A P S T F F K T P A E N E Y L G L V P      1254
Db      634  K T V F E S S F Y M I Q S G N H Q I N D N P O Y Q D F L P N E T K R N G L K V P A A E N E Y L R V A P      689

RESULT 12
TVYUH
protein-tyrosine kinase (EC 2.7.1.112) erbB - avian erythroblastosis virus (strain H)
C:Species: avian erythroblastosis virus
C>Date: 18-Apr-1984 #sequence_revision 18-Apr-1984 #text_change 05-Oct-2004
C:Accession: A00644; A38022
R:Yamamoto, T.; Nishida, T.; Miyajima, N.; Kawai, S.; Ooi, T.; Toyoshima, K.
Cell 35. 71-78, 1983
A>Title: The erbB gene of avian erythroblastosis virus is a member of the src gene famil
A:Reference number: A00644; MUID:84026539; PMID:6313229
A:Accession: A00644
A:Molecule type: DNA
A:Residues: 1-604 <YAM>
A:Cross-references: UNIPROT:P00535; UNIPARC:UPI000028A64; GB:K01216; NID:G209676; PIDN:
Science 224. 1456-1459, 1984
A>Title: Sequencing the erbA gene of avian erythroblastosis virus reveals a new type of
A:Reference number: A38022; MUID:84223957; PMID:6328658
A:Accession: A38022
A:Molecule type: DNA
A:Residues: 1-28, 'W', 30-139, 'F', 141-145, 'V', 147-152 <DBB>
A:Cross-references: UNIPARC:UPI00001725C9; GB:K02006
C:Gene: erbB
C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology

```

C:Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific  
F:130-395/Domain: protein kinase homology <KIN>  
F:138-146/Region: protein kinase ATP-binding motif  
F:165/Active site: Lys #status predicted

Query Match 25.0%; Score 1702; DB 1; Length 604;

Best Local Similarity 52.2%; Pred. No. 3e-64; Mismatches 126; Indels 126; Gaps 16;

Matches 360; Conservative 76; Mismatch 126; Indels 126; Gaps 16;

587 CAHYKDPFCVACRCPGKPLSLVMPKPPDESGACPPCINCHSCVDDDDGCGPMEQ 646

3 CAHFDGHCYKACAGVAGENDTL-VAKYADANAACOLCPNCTRGCKGGLGCP--- 58

647 RASPLTISIAVV-GILLVVLGVVFGILLKROOKIRKRYTMRRLDDETEYHADGGKVP 705

59 NGSTPSTIAAGVGGGLCLVVGIGILYARR-HIVAKRLRLRLQRELVPLTPSGE 117

706 MPNOQMRILKTELKRYKVLGSGAFGVYKGIWPDGENYKIPVAKVLRBENTSPKANK 765

118 APNOAHRLILKETEKVKVVGSGAFGIYKGLWIPBEKRYKIPVAKVLRBENTSPKANK 177

766 EILDEAYVMAVGSPPYVSRLLGICLTSTVOLVTOIMPGCLLDHYRNRRLGSGODLLNW 825

178 EILDEAYVMAVGSPPYVSRLLGICLTSTVOLVTOIMPGCLLDHYRNRRLGSGODLLNW 237

826 CMOJAKGMSYLEDEVRLVARDLAARVVLVKSFNHYKITDPGLARLLDIDETEYHADGGKVP 885

238 CVOJAKGMSYLEDEVRLVARDLAARVVLVKSFNHYKITDPGLARLLDIDETEYHADGGKVP 297

886 IKMMALBSILRRRTTHOSDWSYGVYTWELMTFGAKPYDGIIPAREIPDLLEKGRLLPQP 945

298 IKMMALBSILRRRTTHOSDWSYGVYTWELMTFGAKPYDGIIPAREIPDLLEKGRLLPQP 357

946 ICTDVMIMVCMWIDSECPREFELVSESRMARDPQRFVVIQ-NEIDLPAAPLSTF 1004

358 ICTDVMIMVCMWIDSECPREFELVSESRMARDPQRFVVIQ-NEIDLPAAPLSTF 417

1005 YRSLLDDMDGLVDAEYLVPOQGFPCRPDPAAGVWVHRRHSSSTRSGGADLLGLE 1064

418 YRLTMEEDMEDIDADDEYLVPHQGF-----NSPST----- 449

1065 PSEBAPRSPV-----APSEAGSDVPFGDLGMAKAGLQSLPHNDSPRLORYSEDPV 1119

450 -----SRPLSLSLSATSNNSATNCID-----RNGQGHPRVREDSFVGRYSDDPTGN 495

1120 LPSEFT--DGYVAPLTCSPPQRYVNOQVPRPSPRSGRLPAARPAATLERPRTLSRGK 1177

496 PLEBSIDDFL-----PAREYVNO--LMPKKPSTAM----- 524

1178 NGVAVKDFAF-----GGAVENTEYLTPOGGAAPQHPAPFSPAFD 1218

525 --VONQIYVIFSLTAISKLPMDSRXONSHSTAVDNPEYL-----NTNQSPLAKTVFE 574

1219 NLTYWDDQPRRGAPSTTFKGTPTAENPEY 1248

575 SSPYMIQSGNHQ-----INLDNPDY 594

## RESULT 13

epidermal growth factor receptor - fruit fly (Drosophila melanogaster)

N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbb

C:Species: Drosophila melanogaster

C:Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 05-Oct-2004

C:Accession: A00640; A38021

R:Flymen, E.; Glazer, L.; Segal, D.; Schlessinger, J.; Shilo, B.Z.

Cell 40, 599-607, 1985

A:Title: The Drosophila EGF receptor gene homolog: conservation of both hormone binding

A:Reference number: A00640; MUID:85124611; PMID:2982499

A:Accession: A00640

A:Molecule type: DNA

A:Residues: 1-1330 <LIV>

A:Cross-references: UNIPROT:P04412; UNIPARC:UPI00001725C6; EMBL:K03054

R:Madsworth, S.C.; Vincent II, W.S.; Blodieu-Wentworth, D.

Nature 314, 178-180, 1985

A:Title: A Drosophila genomic sequence with homology to human epidermal growth factor rec

A:Reference number: A38021; MUID:85137938; PMID:2983222

A:Accession: A38021

A:Molecule type: DNA

A:Residues: 'A', 832-866, 'V', 866-943, 'QTPSLVK' <MAD>

A:Cross-references: UNIPARC:UPI0000168BD5; EMBL:X02293; NID:97922; PIDN:CAA26157.1; PID:'

C:Comment: This sequence is tentative because the introns have not been identified.

C:Genetics:

A:Gene: FlyBase:Bgf

A:Cross-references: FlyBase:FBgn0003731

A:Map position: 2 57F

C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho

F:1732/Domain: extracellular #status predicted <EXT>

F:733-764/Domain: transmembrane #status predicted <TM>

F:765-1330/Domain: intracellular #status predicted <INT>

F:808-1072/Domain: protein kinase homology <KIN>

F:816-824/Region: protein kinase ATP-binding motif

F:122,300,324,363,518,688,695,700/Binding site: carbohydrate (Asn) (covalent) #status pr

F:774/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

F:843/Active site: Lys #status predicted

F:1181/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predict

Query Match 24.3%; Score 1653.5; DB 1; Length 1330;

Best Local Similarity 30.0%; Pred. No. 6.9e-62;

Matches 414; Conservative 179; Mismatches 415; Indels 371; Gaps 39;

80 VQGVVLIANQVROVPLQRLRIYRGTLF-----EDNYALALVNDGPNLNTTPTVQASP 134

38 ITNYIVIGLRLIPCTLSYRLQIIRGRITFSLVSEETALFV-----TY 81

135 GGLREIOLRSITELIKGGLVLIORNPOLCYODTILMKDIFPKNNOLATLIDTNRBACHP 194

82 SKMTLEIPLRLVYLVNQVGFHNNYVLCMKRTIOMSVISVNGPDAYVNOFTABERCPK 141

195 CSPWCKSGRCWGSSEDDQSLTRTVACAGCA--RCKGRLPTDCHEOCMAAGTGPKHSDC 252

142 CHSCTHG-CWGBRPKNCKQKFSKLTCSPOCAGRCGRPKPRECHLFCAGGCTPTQKDC 200

253 LACILFHNSGICELHCAALVTYNTDTPESMPNPRGRTFFGASCYTACPYNTLSDVSC 312

201 IACNFFDEAVSKEECPMKRYNTYVLTNPBGKAYAGATCYKECP-GILLRDNACV 259

313 LVCPRLHQBETAEDGTORCEKSCPKCARVCYGLGMEHLREYAVTSANIOEFACCKIFG 372

260 RSCPDQMDKGE-----CYPNGPCPKTCBTVLH-----AGNIDSFRCYVYIDG 306

373 SLAFLEPSFGC--DPASNTA-----PLQEBQLQVFTLEITGYLYISAMPDSLPLSLV 424

307 NIRLIDTDFSGFOVYVNTMGPRYIPLDPERREVFSTVKEITGYLIEGTHPQFRILSY 366

425 FQNLQVIRGRILNGLV-SITLQGLISWGLSRLRSLGGLALIHNTNHLCFHYTPWD 483

367 FRNLETHIGRLQWESMFAALAIYKSLYSLEMRNLKQISSGSVVIQHNRLCYVSNIRWP 426

484 QLFNPHQALLHTANRPEDEC----- 504

427 AIQKEPQKXWVWENLRADLCCKFTLLISVQNNIINHIFALICRKNHLLGSVQRRL 486

505 ----- 504

487 GSWHGVPLVLOELQFWHLHRLMLYQVINSTQDKSNHQLTDACYSVPSTLTIER 546

505 -----VEEGLA-- 510

547 ARVAIAGLAMELEQITTAASAKRHSKTLPAEGROVPRWVFLGVCASAPAGIAPLAGR 606

511 -----CHQLCARGCHGWPGPTQCVNCSQPLRGQCEVECRVLQGLPREVY---NARHCLP 562

607 AVCRKCHPLCELCITNYIHQVCSKTHYGRREGCTEC-----PADHTYDEQRQRCFQ 660



```
Db      58 GLEBGP---NGSKTPSIAGVVGGLCLVVGGLGIGLYLRRR-HIVRRKTLRLLOEREL 113
Qy      697 VEPLTPSGAMPNOAOMRIKETELRKVKYLGSGAFGYKGIWIPDGENVKIPVAIKYLR 756
Db      114 VEPLTPSGEAPNOAHRIKETEFKVKVYLGSGAFGYKGIWIPDGENVKIPVAIKELR 173
Qy      757 ENTSPKANKELIDEAYVNAVGSFVYSRLGI CLTSTVQVLTQMLPYGCLLDHVENRGR 816
Db      174 EATSPKANKELIDEAYVNAVSVDNPHVCRLGI CLTSTVQVLTQMLPYGCLLDYIREHKDN 233
Qy      817 LGSQDLNWCNOIAGMSYLEBVLVHRDLAARNVLVKS PNHVKTITDGLARLDIDETE 876
Db      234 IGSQYLLMWCVOIAKGMNVLBERHLVHRDLAARNVLKTPODVKITDFGLAKOLGADEKE 293
Qy      877 YHADGKVPITKMMALESTILRRRFTHQSDVWSYGVTVWELMTFGAKPYDGI PARZIDLE 936
Db      294 YHABGKVPITKMMALESTILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGI PASEISSYLE 353
Qy      937 KGERLPQPPICTIDVYIMVCKMIDSECRPAFRELVSEFSRMARDPQRFVVIQ-NEDLG 995
Db      354 KGERLPQPPICTIDVYIMVCKMWSADSRPFRELIASFMRARDPRLVLIQGERMH 413
Qy      996 PASPLDSTFYRSILEDMDGDLVDAEYLVPOQGFPCDPAPAGAGMHHRRSSSTRSG 1055
Db      414 LPSPTDSKFYRTLMEEDMEDIVDAEYLVPHQGF-----NSPST--- 454
Qy      1056 GGDLTLGLEPSEBEAPRSPL-----APSEGAGSDVFDGDLGWAAGLQSLPTHDPSPLO 1110
Db      455 -----SRTPLLSSLSATSNNSATNCIDRNG-----H----- 481
Qy      1111 RYSEBDPTVLPSETDGYVAPLTCSPQPEYVNOQDVVPQPPS 1151
Db      482 -----PVREDGFL-----PAPEYVNO--LMPKKPS 504
```

Search completed: January 3, 2006, 11:17:40  
Job time : 59 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: January 3, 2006, 11:17:49 ; Search time 163 Seconds

(without alignments)  
2865.442 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 MELALCRWGLLLALLPPGA.....TFKGTPTAENPEYLGLDVPV 1255

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6815	100.0	1255	3	US-09-765-973-2
2	6815	100.0	1255	3	US-09-854-356-1
3	6815	100.0	1255	3	US-09-930-125-2
4	6815	100.0	1255	3	US-09-441-411-6
5	6815	100.0	1255	4	US-10-207-655-45
6	6815	100.0	1255	4	US-10-313-644-2
7	6815	100.0	1255	4	US-10-418-027-3
8	6815	100.0	1255	4	US-10-394-322A-17
9	6815	100.0	1255	4	US-10-245-871-553
10	6815	100.0	1255	4	US-10-469-162-3
11	6815	100.0	1255	4	US-10-253-286-553
12	6815	100.0	1255	5	US-10-762-128-6
13	6815	100.0	1255	5	US-10-484-067-1
14	6815	100.0	1255	5	US-10-723-860-9
15	6815	100.0	1255	5	US-10-871-708-9
16	6815	100.0	1255	5	US-10-983-340-12
17	6815	100.0	1255	6	US-11-037-713-12
18	6806	99.9	1255	3	US-09-811-123-9
19	6806	99.9	1255	3	US-09-811-115-3
20	6806	99.9	1255	3	US-09-354-533-68
21	6806	99.9	1255	4	US-09-984-092-4
22	6806	99.9	1255	4	US-10-177-293-126
23	6806	99.9	1255	4	US-10-207-498-6
24	6806	99.9	1255	4	US-10-338-730-2
25	6806	99.9	1255	4	US-10-322-892-4
26	6806	99.9	1255	4	US-10-272-437A-28
27	6806	99.9	1255	4	US-10-117-937-594

28	6806	99.9	1255	4	US-10-435-696-36	Sequence 36, Appl
29	6806	99.9	1255	4	US-10-647-005-68	Sequence 68, Appl
30	6806	99.9	1255	4	US-10-441-779C-4	Sequence 4, Appl
31	6806	99.9	1255	4	US-10-734-564-126	Sequence 126, App
32	6806	99.9	1255	4	US-10-657-022-90	Sequence 90, Appl
33	6806	99.9	1255	5	US-10-615-343-17	Sequence 17, Appl
34	6806	99.9	1255	5	US-10-794-514A-1	Sequence 1, Appl
35	6806	99.9	1255	5	US-10-846-113A-27	Sequence 27, Appl
36	6806	99.9	1255	6	US-11-067-064-594	Sequence 594, App
37	6806	99.9	1255	6	US-11-121-347-68	Sequence 68, Appl
38	6806	99.9	1255	6	US-11-067-159-594	Sequence 594, App
39	6804	99.8	1253	4	US-10-146-473-72	Sequence 72, Appl
40	6803	99.8	1255	4	US-10-149-138-4641	Sequence 4641, Ap
41	6803	99.8	1255	4	US-10-149-138-4641	Sequence 4641, Ap
42	6776.5	99.4	1294	5	US-10-956-373-10	Sequence 10, Appl
43	6757	99.1	1255	3	US-09-765-508-2	Sequence 2, Appl
44	6003	88.1	1257	5	US-10-484-067-2	Sequence 2, Appl
45	6001.5	88.1	1256	3	US-09-854-356-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1									
US-09-765-973-2									
; Sequence 2, Application US/09765973									
; Publication No. US20020039573A1									
; GENERAL INFORMATION:									
; APPLICANT: Cheever, Martin A.									
; TITLE OF INVENTION: Hand-Zimmerman, Susan									
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR PREVENTION AND									
; CURRENT FILING DATE: 2001-01-19									
; CURRENT APPLICATION NUMBER: US/09/765,973									
; NUMBER OF SEQ ID NOS: 4									
; SOFTWARE: FastSeq for Windows Version 3.0									
; SEQ ID NO 2									
; LENGTH: 1255									
; TYPR: PRT									
; ORGANISM: Homo sapien									
US-09-765-973-2									
Query Match									
100.0%; Score 6815; DB 3; Length 1255;									
Query Local Similarity 100.0%; Pred. No. 0;									
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MELALCRWGLLLALLPPGA	STVCTGTDKRLPAS	PETHLDMRLHYGCGVVGNTL	60				
DB	1	MELALCRWGLLLALLPPGA	STVCTGTDKRLPAS	PETHLDMRLHYGCGVVGNTL	60				
QY	61	ELTYPTNASSFLDDIOEV	GVYLIANNQVQVPLQRL	RYRGTLFEDNYALAVLNG	120				
DB	61	ELTYPTNASSFLDDIOEV	GVYLIANNQVQVPLQRL	RYRGTLFEDNYALAVLNG	120				
QY	121	DPINNTVTGASPGGLRE	LOBSLTILKGVLIQRNP	OLCYOPTIMKQIFHNQOLA	180				
DB	121	DPINNTVTGASPGGLRE	LOBSLTILKGVLIQRNP	OLCYOPTIMKQIFHNQOLA	180				
QY	181	LTLIDTNRSRACHPCSP	CKSRGWSBQSLTRTV	CAGCARCKGPLETDCHEQC	240				
DB	181	LTLIDTNRSRACHPCSP	CKSRGWSBQSLTRTV	CAGCARCKGPLETDCHEQC	240				
QY	241	AAGCTGPRGSDCLALH	NHSGICELHPALVY	NTDTFESMPNBEGRYTRGASCVTACP	300				
DB	241	AAGCTGPRGSDCLALH	NHSGICELHPALVY	NTDTFESMPNBEGRYTRGASCVTACP	300				
QY	301	YNYLSTDVGSCTLVCP	LHNOEYTAEDGTOR	CEKSKPCARVCYGLGMEHLREVRATVSAN	360				
DB	301	YNYLSTDVGSCTLVCP	LHNOEYTAEDGTOR	CEKSKPCARVCYGLGMEHLREVRATVSAN	360				
QY	361	IOEFAGCKKIRGSLA	FLPESFDGDASTAP	LQPEQLQVFETLBEITLYISAMPDLP	420				
DB	361	IOEFAGCKKIRGSLA	FLPESFDGDASTAP	LQPEQLQVFETLBEITLYISAMPDLP	420				

```
Db 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPQOVFETLEITGYLISAMPDLSL 420
Qy 421 DLSVFONLQVIRGRILHNGAYSLTLQGLISWLGSRILRELGSGALIHNTHLCEVATV 480
Db 421 DLSVFONLQVIRGRILHNGAYSLTLQGLISWLGSRILRELGSGALIHNTHLCEVATV 480
Qy 481 FMDQLFRNPQALHTANRPEDECVGEGLAHQLCARHGWGPPTQCVCNCSQFLRGEC 540
Db 481 FMDQLFRNPQALHTANRPEDECVGEGLAHQLCARHGWGPPTQCVCNCSQFLRGEC 540
Qy 541 VEECVIQLGPREYVNAHCLPCHPECOPONGSVTCFGEADQCACAHYDPPFCVARC 600
Db 541 VEECVIQLGPREYVNAHCLPCHPECOPONGSVTCFGEADQCACAHYDPPFCVARC 600
Qy 601 PSQVPPDLSYMPIMKFPDEBGACOPCPINCHSCVDLDDKCPABORASPLTISIAYVG 660
Db 601 PSQVPPDLSYMPIMKFPDEBGACOPCPINCHSCVDLDDKCPABORASPLTISIAYVG 660
Qy 661 ILVVVLGVFGLIKRRQOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLIKETEL 720
Db 661 ILVVVLGVFGLIKRRQOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLIKETEL 720
Qy 721 RKVVLGSGAGRTYKGIWIPGENVKIPVAIKVIRENTSPKANKELIDELAVMAVGSP 780
Db 721 RKVVLGSGAGRTYKGIWIPGENVKIPVAIKVIRENTSPKANKELIDELAVMAVGSP 780
Qy 781 YVSRLLGI CLTSTVQLTQMPYGLDHRVRENGRLSODLLNCKMOIAKMSYLEDVR 840
Db 781 YVSRLLGI CLTSTVQLTQMPYGLDHRVRENGRLSODLLNCKMOIAKMSYLEDVR 840
Qy 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETRYHADGGKVPDKMMALESILRRFT 900
Db 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETRYHADGGKVPDKMMALESILRRFT 900
Qy 901 HOSDWSGVTYWEIMLTGAKPYDGI PAKEI PDLLEKESRLPQPICTIDYVMIVKMM 960
Db 901 HOSDWSGVTYWEIMLTGAKPYDGI PAKEI PDLLEKESRLPQPICTIDYVMIVKMM 960
Qy 961 IDSECRPFRELVSFSSMARDPQRFVYIQNEDLPASPLDSTFYSRLLEDMDKDLVDA 1020
Db 961 IDSECRPFRELVSFSSMARDPQRFVYIQNEDLPASPLDSTFYSRLLEDMDKDLVDA 1020
Qy 1021 EBYLVPOQGFPCPDPAFGAGVHHRRHSSSTRSGGDLITGLPESSEBAPRSLAPSEG 1080
Db 1021 EBYLVPOQGFPCPDPAFGAGVHHRRHSSSTRSGGDLITGLPESSEBAPRSLAPSEG 1080
Qy 1081 AGSDVPDDDLGMAKGLQSLPTHDPSLQRYSEDPVPLBSETGVAVPLTCSQPBEYV 1140
Db 1081 AGSDVPDDDLGMAKGLQSLPTHDPSLQRYSEDPVPLBSETGVAVPLTCSQPBEYV 1140
Qy 1141 NOPDVRPOPSPREBGLPARAPAGATLERPKTLSPGKGVVADVAFGGAVENTPEYLTPO 1200
Db 1141 NOPDVRPOPSPREBGLPARAPAGATLERPKTLSPGKGVVADVAFGGAVENTPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPAFDNLVYWDODPBERGAPSTFKGTPTAENPEYLGIDVPV 1255
Db 1201 GGAAPQHPHPAPAFDNLVYWDODPBERGAPSTFKGTPTAENPEYLGIDVPV 1255

RESULT 2
US-09-854-356-1
; Sequence 1, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheyssen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
```

```
Query Match 100.0%; Score 6815; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-854-356-1
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: DOMAIN
; LOCATION: (1)..(653)
; OTHER INFORMATION: extracellular domain (ECD)
; NAME/KEY: DOMAIN
; LOCATION: (676)..(1255)
; OTHER INFORMATION: intracellular domain (ICD)
; NAME/KEY: DOMAIN
; LOCATION: (990)..(1255)
; OTHER INFORMATION: phosphorylation domain (PD)
; NAME/KEY: DOMAIN
; LOCATION: (990)..(1048)
; OTHER INFORMATION: fragment of the phosphorylation domain, preferred
; OTHER INFORMATION: portion (delta PD)
; US-09-854-356-1

Qy 1 MELALCRWGLLALLPPGAASSTOVCTGDMKRLPASPEYHLMRLHYGCGVOGNTL 60
Db 1 MELALCRWGLLALLPPGAASSTOVCTGDMKRLPASPEYHLMRLHYGCGVOGNTL 60
Qy 61 ELTLPNMAISFLQDIQEVGVYLIHNOVROYPLRLRIVRGTOLEPBNYALAVLDNG 120
Db 61 ELTLPNMAISFLQDIQEVGVYLIHNOVROYPLRLRIVRGTOLEPBNYALAVLDNG 120
Qy 121 DPLNNTTVPVTSAGSGRLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDI FHKNNQLA 180
Db 121 DPLNNTTVPVTSAGSGRLRELQRLSLTEILKGVLIQRPOLCYQDTILMKDI FHKNNQLA 180
Qy 181 LTLIDTNRSRACHPCSPMCKSGRCWGSSEDCQSLTRTVCAAGCARCKGRLPTDCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMCKSGRCWGSSEDCQSLTRTVCAAGCARCKGRLPTDCHEQC 240
Qy 241 AAGCTGPHSHCLCLPHNHSIGICELHCPALVTYNTDTPESMPNDEGRYTFGASCVTACP 300
Db 241 AAGCTGPHSHCLCLPHNHSIGICELHCPALVTYNTDTPESMPNDEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLYCPLHNDVTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
Db 301 YNYLSTDVGSCTLYCPLHNDVTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
Qy 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPQOVFETLEITGYLISAMPDLSL 420
Db 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPQOVFETLEITGYLISAMPDLSL 420
Qy 421 DLSVFONLQVIRGRILHNGAYSLTLQGLISWLGSRILRELGSGALIHNTHLCEVATV 480
Db 421 DLSVFONLQVIRGRILHNGAYSLTLQGLISWLGSRILRELGSGALIHNTHLCEVATV 480
Qy 481 FMDQLFRNPQALHTANRPEDECVGEGLAHQLCARHGWGPPTQCVCNCSQFLRGEC 540
Db 481 FMDQLFRNPQALHTANRPEDECVGEGLAHQLCARHGWGPPTQCVCNCSQFLRGEC 540
Qy 541 VEECVIQLGPREYVNAHCLPCHPECOPONGSVTCFGEADQCACAHYDPPFCVARC 600
Db 541 VEECVIQLGPREYVNAHCLPCHPECOPONGSVTCFGEADQCACAHYDPPFCVARC 600
Qy 601 PSQVPPDLSYMPIMKFPDEBGACOPCPINCHSCVDLDDKCPABORASPLTISIAYVG 660
Db 601 PSQVPPDLSYMPIMKFPDEBGACOPCPINCHSCVDLDDKCPABORASPLTISIAYVG 660
```

QY 661 ILVVVLGVFGILLKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETEL 720  
DB 661 ILVVVLGVFGILLKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETEL 720  
QY 721 RKVKVLGSGAGFTYVKGIMI PDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAVGSP 780  
DB 721 RKVKVLGSGAGFTYVKGIMI PDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAVGSP 780  
QY 781 YVSRLLGICLTSTVQLVQLMPYGCILDHVENRGRLSQDILNMCQIAKMSYLEDR 840  
DB 781 YVSRLLGICLTSTVQLVQLMPYGCILDHVENRGRLSQDILNMCQIAKMSYLEDR 840  
QY 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETRYHADGKVPKIMMALESILRRFT 900  
DB 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETRYHADGKVPKIMMALESILRRFT 900  
QY 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKGERLPQPICTIDVYMIWCKM 960  
DB 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKGERLPQPICTIDVYMIWCKM 960  
QY 961 IDSECRPRRELVSERSSMAADPQRFVYIQNEDLGAPSLDSTFYRSLLEDDMGDLVDA 1020  
DB 961 IDSECRPRRELVSERSSMAADPQRFVYIQNEDLGAPSLDSTFYRSLLEDDMGDLVDA 1020  
QY 1021 BEYLVPOQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBPSEBAPRSPLAPSEG 1080  
DB 1021 BEYLVPOQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBPSEBAPRSPLAPSEG 1080  
QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
DB 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
QY 1141 NOPDVRPOPSPREGPLPAAPAGATLERPKTSLSGKGVKQVAVAFGAVENPEYLTPO 1200  
DB 1141 NOPDVRPOPSPREGPLPAAPAGATLERPKTSLSGKGVKQVAVAFGAVENPEYLTPO 1200  
QY 1201 GGAAPQPPPPAFSPAFDNLVYMDODPPERGA PPSFTFKGTPTAENPEYLGDLVPV 1255  
DB 1201 GGAAPQPPPPAFSPAFDNLVYMDODPPERGA PPSFTFKGTPTAENPEYLGDLVPV 1255

RESULT 3  
US-09-930-125-2  
; Sequence 2, Application US/09930125  
; Publication No. US20020193329A1  
; GENERAL INFORMATION:  
; APPLICANT: Hand-Zimmerman, Susan  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Lodge, Michael J.  
; APPLICANT: Kates, Michael D.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Vedvick, Thomas S.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
; TITLE OF INVENTION: OF HBR-2/NEU-ASSOCIATED MALIGNANCIES  
; FILE REFERENCE: 210121.544  
; CURRENT APPLICATION NUMBER: US/09/930,125  
; CURRENT FILING DATE: 2001-08-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-930-125-2

Query Match 100.0%; Score 6815; DB 3; Length 1255;  
Beet Local Similarity 100.0%; Pred. No. 0;  
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALCRWGLLALLPFGAASVQCTGTDMKRLRPASBETHLDMRLHLYGCGQVVOGNTL 60  
DB 1 METALCRWGLLALLPFGAASVQCTGTDMKRLRPASBETHLDMRLHLYGCGQVVOGNTL 60

DB 1 METALCRWGLLALLPFGAASVQCTGTDMKRLRPASBETHLDMRLHLYGCGQVVOGNTL 60  
QY 61 BLTYLPTNASISFLDIOEVQGYVLIAHQVQVPELQRLRIYRGTLFEDNYALAVLDNG 120  
DB 61 BLTYLPTNASISFLDIOEVQGYVLIAHQVQVPELQRLRIYRGTLFEDNYALAVLDNG 120  
QY 121 DPLANTTPTVSGAPGLREBLQRLSLEILKGGVLIQRNPOLCYQDTILMKDI FHKNNQIA 180  
DB 121 DPLANTTPTVSGAPGLREBLQRLSLEILKGGVLIQRNPOLCYQDTILMKDI FHKNNQIA 180  
QY 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSILTRYCAGGCAKCGPLPTDCBEC 240  
DB 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSILTRYCAGGCAKCGPLPTDCBEC 240  
QY 241 AAGCTGPHNSDCLAHFNHSGICLHCPALVTYNTDFESPMNPEGRTFGASCVTYACP 300  
DB 241 AAGCTGPHNSDCLAHFNHSGICLHCPALVTYNTDFESPMNPEGRTFGASCVTYACP 300  
QY 301 YNYLSTDVGSCTLVCPLNQEVTAEDGTORCEKSKPCARVCYGLAMEHLREVRATVSAN 360  
DB 301 YNYLSTDVGSCTLVCPLNQEVTAEDGTORCEKSKPCARVCYGLAMEHLREVRATVSAN 360  
QY 361 IOEFAGCKKIFGSLAFPLPSFPDGPASNTAPLOPEQLOVFTLEBITGYLISAMPDSL 420  
DB 361 IOEFAGCKKIFGSLAFPLPSFPDGPASNTAPLOPEQLOVFTLEBITGYLISAMPDSL 420  
QY 421 DLSVFQNTQVIRGRILHNAGVSLTLOGISWGLRSLEIGSSGLAIHNNHLCFVHTV 480  
DB 421 DLSVFQNTQVIRGRILHNAGVSLTLOGISWGLRSLEIGSSGLAIHNNHLCFVHTV 480  
QY 481 PWDQLFRPHQALLHTANRPEDECEVGEGLACHQLCARGHCWGPPTQCNCQSLRQDEC 540  
DB 481 PWDQLFRPHQALLHTANRPEDECEVGEGLACHQLCARGHCWGPPTQCNCQSLRQDEC 540  
QY 541 VEECKVLOGLPREYVNAHCLPCHPECOPQNSVTCFPEADQCAVACHYKOPPCVAVRC 600  
DB 541 VEECKVLOGLPREYVNAHCLPCHPECOPQNSVTCFPEADQCAVACHYKOPPCVAVRC 600  
QY 601 PSGVPELSTYMPIMKFPDEBGACOPCINCHSCVDLDDKGPAPORASPLTISIAYVG 660  
DB 601 PSGVPELSTYMPIMKFPDEBGACOPCINCHSCVDLDDKGPAPORASPLTISIAYVG 660  
QY 661 ILVVVLGVFGILLKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETEL 720  
DB 661 ILVVVLGVFGILLKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETEL 720  
QY 721 RKVKVLGSGAGFTYVKGIMI PDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAVGSP 780  
DB 721 RKVKVLGSGAGFTYVKGIMI PDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAVGSP 780  
QY 781 YVSRLLGICLTSTVQLVQLMPYGCILDHVENRGRLSQDILNMCQIAKMSYLEDR 840  
DB 781 YVSRLLGICLTSTVQLVQLMPYGCILDHVENRGRLSQDILNMCQIAKMSYLEDR 840  
QY 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETRYHADGKVPKIMMALESILRRFT 900  
DB 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETRYHADGKVPKIMMALESILRRFT 900  
QY 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKGERLPQPICTIDVYMIWCKM 960  
DB 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKGERLPQPICTIDVYMIWCKM 960  
QY 961 IDSECRPRRELVSERSSMAADPQRFVYIQNEDLGAPSLDSTFYRSLLEDDMGDLVDA 1020  
DB 961 IDSECRPRRELVSERSSMAADPQRFVYIQNEDLGAPSLDSTFYRSLLEDDMGDLVDA 1020  
QY 1021 BEYLVPOQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBPSEBAPRSPLAPSEG 1080  
DB 1021 BEYLVPOQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBPSEBAPRSPLAPSEG 1080  
QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
DB 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140

```
Oy 1141 NOPVRRPQPSRREGPLPAARPAAGATLERPKTILSPGKGVVMDVFAFGAVENBEYLTPQ 1200
Db 1141 NOPVRRPQPSRREGPLPAARPAAGATLERPKTILSPGKGVVMDVFAFGAVENBEYLTPQ 1200
Oy 1201 GGAAPQHPHPAFSFAFDNLVYWDQDPPERGAPSTFKGTPTTANPEYLGIDVY 1255
Db 1201 GGAAPQHPHPAFSFAFDNLVYWDQDPPERGAPSTFKGTPTTANPEYLGIDVY 1255

RESULT 4
US-09-441-411-6
; Sequence 6, Application US/09441411
; Publication No. US2003008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disib, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 73003.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-411-6

Query Match 100.0%; Score 6815; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MELAALCRWGLLALLPPGAASSTOVCTGDMKRLPASPEETHLDMRLHLYGSCVYQGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTOVCTGDMKRLPASPEETHLDMRLHLYGSCVYQGNL 60
Oy 61 ELTYLPTNASLSFODDIOEVQGVYLAHNQVQVPLQNRVGRVGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFODDIOEVQGVYLAHNQVQVPLQNRVGRVGTQLFEDNYALAVLDNG 120
Oy 121 DPLANTTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
Db 121 DPLANTTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
Oy 181 LTLIDTNRSRACHPCSPWCKSGRSGESSEDCQSLTRTVCAAGGCARCKGPLEPTDCHEQC 240
Db 181 LTLIDTNRSRACHPCSPWCKSGRSGESSEDCQSLTRTVCAAGGCARCKGPLEPTDCHEQC 240
Oy 241 AAGCGPRHSDCLALHFNHSGICELHCPALVTYNTDFESPNNEGRYTGASCVTACP 300
Db 241 AAGCGPRHSDCLALHFNHSGICELHCPALVTYNTDFESPNNEGRYTGASCVTACP 300
Oy 301 YNYLSTDVGSCTLVPLHNOEVTADGTQRCEKSKPCARVYGLGMEHLREVRVTSAN 360
Db 301 YNYLSTDVGSCTLVPLHNOEVTADGTQRCEKSKPCARVYGLGMEHLREVRVTSAN 360
Oy 361 IQEPAGCKKIFGSLAFLPESFGDPASTAPLQAPQOLQFFETLEBITGTYLSAMPDSL 420
Db 361 IQEPAGCKKIFGSLAFLPESFGDPASTAPLQAPQOLQFFETLEBITGTYLSAMPDSL 420
Oy 421 DLVSFQNLQVLRGRILHNGAVSLTLOGLISWLGRSLRELGSGALIHNNTHLCFVHTV 480
Db 421 DLVSFQNLQVLRGRILHNGAVSLTLOGLISWLGRSLRELGSGALIHNNTHLCFVHTV 480
Oy 481 PMDQFLRPHQALHTANRPEDECVGEGLACHQLCARGHCMWPGPTQCVNCSQFLRGQEC 540
Db 481 PMDQFLRPHQALHTANRPEDECVGEGLACHQLCARGHCMWPGPTQCVNCSQFLRGQEC 540
Oy 541 VEECRVLOGLPREYVNAHRCLECHPECOFQNGSVTCFGEADQVCAAHYKDPFCVARC 600
Db 541 VEECRVLOGLPREYVNAHRCLECHPECOFQNGSVTCFGEADQVCAAHYKDPFCVARC 600
```

```
Db 541 VEECRVLOGLPREYVNAHRCLECHPECOFQNGSVTCFGEADQVCAAHYKDPFCVARC 600
Oy 601 PSQVPLSLYMPILMKFDEEGACQPCPINCTHSCVDLDDKGCAPQASPLTISIAVVG 660
Db 601 PSQVPLSLYMPILMKFDEEGACQPCPINCTHSCVDLDDKGCAPQASPLTISIAVVG 660
Oy 661 ILVVLGVVFGIILKRRQOKIRKTYMRLLQETELVEPLTPSGAMRYOQMRILKXTEL 720
Db 661 ILVVLGVVFGIILKRRQOKIRKTYMRLLQETELVEPLTPSGAMRYOQMRILKXTEL 720
Oy 721 RKVVLGSGAFGYTKGIWI PDGENYKIPVAIKYLRNTSPKAKELIDEAYVAGVQSP 780
Db 721 RKVVLGSGAFGYTKGIWI PDGENYKIPVAIKYLRNTSPKAKELIDEAYVAGVQSP 780
Oy 781 YVSRLLGICLTSITVQVLTQMLPYGCLLDHYRENGRLSGODLNMCMQIAKMSYLEDVR 840
Db 781 YVSRLLGICLTSITVQVLTQMLPYGCLLDHYRENGRLSGODLNMCMQIAKMSYLEDVR 840
Oy 841 LVHRDLAARVYLKSPHNVKITDFGLARLDDIDETBYHADGKVPIKMALESLIRRF 900
Db 841 LVHRDLAARVYLKSPHNVKITDFGLARLDDIDETBYHADGKVPIKMALESLIRRF 900
Oy 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRVLPQPICTIDVYIMVCKM 960
Db 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRVLPQPICTIDVYIMVCKM 960
Oy 961 IDSECRPRFRELVESEFARMADPQRFVYIIONEDIGPASPLDSTFYRSLLEDMDGLVDA 1020
Db 961 IDSECRPRFRELVESEFARMADPQRFVYIIONEDIGPASPLDSTFYRSLLEDMDGLVDA 1020
Oy 1021 EBYLVPOQGFCDPDPAPGAGMVHHRSSSTRSGGDLTLGLEPSEBEARSLAASEG 1080
Db 1021 EBYLVPOQGFCDPDPAPGAGMVHHRSSSTRSGGDLTLGLEPSEBEARSLAASEG 1080
Oy 1081 AGSDVPFGDLMGMAKGLQSLPTHDSPLQRYSDPVPPLPSETDGVAPLTCSPQREYV 1140
Db 1081 AGSDVPFGDLMGMAKGLQSLPTHDSPLQRYSDPVPPLPSETDGVAPLTCSPQREYV 1140
Oy 1141 NOPVRRPQPSRREGPLPAARPAAGATLERPKTILSPGKGVVMDVFAFGAVENBEYLTPQ 1200
Db 1141 NOPVRRPQPSRREGPLPAARPAAGATLERPKTILSPGKGVVMDVFAFGAVENBEYLTPQ 1200
Oy 1201 GGAAPQHPHPAFSFAFDNLVYWDQDPPERGAPSTFKGTPTTANPEYLGIDVY 1255
Db 1201 GGAAPQHPHPAFSFAFDNLVYWDQDPPERGAPSTFKGTPTTANPEYLGIDVY 1255

RESULT 5
US-10-207-655-45
; Sequence 45, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-45

Query Match 100.0%; Score 6815; DB 4; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MELAALCRWGLLALLPPGAASSTOVCTGDMKRLPASPEETHLDMRLHLYGSCVYQGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTOVCTGDMKRLPASPEETHLDMRLHLYGSCVYQGNL 60
```



61 ELTYLPTNASISFLODIOEVQGYVLIANQVROVPLQRLRIYRGTOLEFEDNYALAVLDNG 120  
61 ELTYLPTNASISFLODIOEVQGYVLIANQVROVPLQRLRIYRGTOLEFEDNYALAVLDNG 120  
QY 121 DPLANTTPTVGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNQIA 180  
121 DPLANTTPTVGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNQIA 180  
Db 121 DPLANTTPTVGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNQIA 180  
QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEBDCQSLTRTYCAGGCARCKPLPTDCCHQC 240  
181 LTLIDTNSRACHPCSPMKSGRCWGESSEBDCQSLTRTYCAGGCARCKPLPTDCCHQC 240  
Db 181 LTLIDTNSRACHPCSPMKSGRCWGESSEBDCQSLTRTYCAGGCARCKPLPTDCCHQC 240  
QY 241 AAGCTGPRHSDCLACHFNHSGICELHCPALVTYNTDFESMPNREGRTTFGASCVTACP 300  
241 AAGCTGPRHSDCLACHFNHSGICELHCPALVTYNTDFESMPNREGRTTFGASCVTACP 300  
Db 241 AAGCTGPRHSDCLACHFNHSGICELHCPALVTYNTDFESMPNREGRTTFGASCVTACP 300  
QY 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360  
301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360  
Db 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360  
QY 361 IOEPFAGCKKI FGSLLAFPLPESFDGDPASNTAPLOPBOLOVFTLEBEITGYLISAMPDSL 420  
361 IOEPFAGCKKI FGSLLAFPLPESFDGDPASNTAPLOPBOLOVFTLEBEITGYLISAMPDSL 420  
Db 361 IOEPFAGCKKI FGSLLAFPLPESFDGDPASNTAPLOPBOLOVFTLEBEITGYLISAMPDSL 420  
QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGIGISWLGIRSLRELSGLALIHNTHLCFVHTV 480  
421 DLSVFQNLQVIRGRILHNGAYSLTLOGIGISWLGIRSLRELSGLALIHNTHLCFVHTV 480  
Db 421 DLSVFQNLQVIRGRILHNGAYSLTLOGIGISWLGIRSLRELSGLALIHNTHLCFVHTV 480  
QY 481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540  
481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540  
Db 481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540  
QY 541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600  
541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600  
Db 541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600  
QY 601 PSGVPRDLSYMPIMKFPDEEGACOPCPICHTSCVDLDDKGPABORASPLSIISAVYG 660  
601 PSGVPRDLSYMPIMKFPDEEGACOPCPICHTSCVDLDDKGPABORASPLSIISAVYG 660  
Db 601 PSGVPRDLSYMPIMKFPDEEGACOPCPICHTSCVDLDDKGPABORASPLSIISAVYG 660  
QY 661 ILLVVVLGVFGLILIKRQOKIRKTYTMRILQETELVEPLTPSGAMPNOQRILKETEL 720  
661 ILLVVVLGVFGLILIKRQOKIRKTYTMRILQETELVEPLTPSGAMPNOQRILKETEL 720  
Db 661 ILLVVVLGVFGLILIKRQOKIRKTYTMRILQETELVEPLTPSGAMPNOQRILKETEL 720  
QY 721 RKVYLVGSAFGTYKGIWIPDGENVKIPVAIKVIRENTSPYANKIILDEAYVMAGVGP 780  
721 RKVYLVGSAFGTYKGIWIPDGENVKIPVAIKVIRENTSPYANKIILDEAYVMAGVGP 780  
Db 721 RKVYLVGSAFGTYKGIWIPDGENVKIPVAIKVIRENTSPYANKIILDEAYVMAGVGP 780  
QY 781 YVSRLLGICLTSTVOLVQOLMPYGCILDHVENRGRISQDILLNCMQIAKMSYLEDR 840  
781 YVSRLLGICLTSTVOLVQOLMPYGCILDHVENRGRISQDILLNCMQIAKMSYLEDR 840  
Db 781 YVSRLLGICLTSTVOLVQOLMPYGCILDHVENRGRISQDILLNCMQIAKMSYLEDR 840  
QY 841 LVHRDLAARNVAVKSPNHVKITDPGLARLLDIDETEHADGGKVPDKMMALESIILRRRT 900  
841 LVHRDLAARNVAVKSPNHVKITDPGLARLLDIDETEHADGGKVPDKMMALESIILRRRT 900  
Db 841 LVHRDLAARNVAVKSPNHVKITDPGLARLLDIDETEHADGGKVPDKMMALESIILRRRT 900  
QY 901 HOSDWSVSGVTWELMTFGAKPYDGI PAEIPDLLEKGERLPQPICTIDVTMINKCM 960  
901 HOSDWSVSGVTWELMTFGAKPYDGI PAEIPDLLEKGERLPQPICTIDVTMINKCM 960  
Db 901 HOSDWSVSGVTWELMTFGAKPYDGI PAEIPDLLEKGERLPQPICTIDVTMINKCM 960  
QY 961 IDSECRPRRELVSERPMARDPQRFVVIQNEIDLGPASPLDSTFKSLLEDDMDGLVDA 1020  
961 IDSECRPRRELVSERPMARDPQRFVVIQNEIDLGPASPLDSTFKSLLEDDMDGLVDA 1020  
Db 961 IDSECRPRRELVSERPMARDPQRFVVIQNEIDLGPASPLDSTFKSLLEDDMDGLVDA 1020  
QY 1021 BEYLVPQOGFCPPDPAFGAGVNHHRSSSTRSGGDLTLGLBSEEBEAPRSLAPSEG 1080  
1021 BEYLVPQOGFCPPDPAFGAGVNHHRSSSTRSGGDLTLGLBSEEBEAPRSLAPSEG 1080  
Db 1021 BEYLVPQOGFCPPDPAFGAGVNHHRSSSTRSGGDLTLGLBSEEBEAPRSLAPSEG 1080  
QY 1081 AGSDVFDGLGMAKAGLOSLPTHDBPLQRYSEBDPTVPLPSETDGYVALPLTCSPOPEY 1140  
1081 AGSDVFDGLGMAKAGLOSLPTHDBPLQRYSEBDPTVPLPSETDGYVALPLTCSPOPEY 1140  
Db 1081 AGSDVFDGLGMAKAGLOSLPTHDBPLQRYSEBDPTVPLPSETDGYVALPLTCSPOPEY 1140

QY 1141 NOPDVRPOPSPBREGPLPAAPAGATLTERPKTLPBGKGVKVDVAFGGAIVENPEYLTPO 1200  
1141 NOPDVRPOPSPBREGPLPAAPAGATLTERPKTLPBGKGVKVDVAFGGAIVENPEYLTPO 1200  
Db 1141 NOPDVRPOPSPBREGPLPAAPAGATLTERPKTLPBGKGVKVDVAFGGAIVENPEYLTPO 1200  
QY 1201 GGAAPQHPPPAFSAFBNLYWDDPPERGAPESTFEGTPTAENPEYLGIDVP 1255  
1201 GGAAPQHPPPAFSAFBNLYWDDPPERGAPESTFEGTPTAENPEYLGIDVP 1255  
Db 1201 GGAAPQHPPPAFSAFBNLYWDDPPERGAPESTFEGTPTAENPEYLGIDVP 1255

RESULT 6  
US-10-313-644-2  
; Sequence 2, Application US/10313644  
; Publication No. US20030157119A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Hand-Zimmerman, Susan  
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL  
; FILE REFERENCE: 210121.483C3  
; CURRENT APPLICATION NUMBER: US/10/313,644  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1255  
; TYPE: PRF  
; ORGANISM: Homo sapien  
US-10-313-644-2

Query Match 100.0%; Score 6815; DB 4; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELALCWGGLLALLPFGAASVOYCTGTDMKLRIPASBETHLDMRLHYYGCVYQGNL 60  
1 MELALCWGGLLALLPFGAASVOYCTGTDMKLRIPASBETHLDMRLHYYGCVYQGNL 60  
Db 1 MELALCWGGLLALLPFGAASVOYCTGTDMKLRIPASBETHLDMRLHYYGCVYQGNL 60  
QY 61 ELTYLPTNASISFLODIOEVQGYVLIANQVROVPLQRLRIYRGTOLEFEDNYALAVLDNG 120  
61 ELTYLPTNASISFLODIOEVQGYVLIANQVROVPLQRLRIYRGTOLEFEDNYALAVLDNG 120  
Db 61 ELTYLPTNASISFLODIOEVQGYVLIANQVROVPLQRLRIYRGTOLEFEDNYALAVLDNG 120  
QY 121 DPLANTTPTVGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNQIA 180  
121 DPLANTTPTVGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNQIA 180  
Db 121 DPLANTTPTVGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNQIA 180  
QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEBDCQSLTRTYCAGGCARCKPLPTDCCHQC 240  
181 LTLIDTNSRACHPCSPMKSGRCWGESSEBDCQSLTRTYCAGGCARCKPLPTDCCHQC 240  
Db 181 LTLIDTNSRACHPCSPMKSGRCWGESSEBDCQSLTRTYCAGGCARCKPLPTDCCHQC 240  
QY 241 AAGCTGPRHSDCLACHFNHSGICELHCPALVTYNTDFESMPNREGRTTFGASCVTACP 300  
241 AAGCTGPRHSDCLACHFNHSGICELHCPALVTYNTDFESMPNREGRTTFGASCVTACP 300  
Db 241 AAGCTGPRHSDCLACHFNHSGICELHCPALVTYNTDFESMPNREGRTTFGASCVTACP 300  
QY 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360  
301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360  
Db 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360  
QY 361 IOEPFAGCKKI FGSLLAFPLPESFDGDPASNTAPLOPBOLOVFTLEBEITGYLISAMPDSL 420  
361 IOEPFAGCKKI FGSLLAFPLPESFDGDPASNTAPLOPBOLOVFTLEBEITGYLISAMPDSL 420  
Db 361 IOEPFAGCKKI FGSLLAFPLPESFDGDPASNTAPLOPBOLOVFTLEBEITGYLISAMPDSL 420  
QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGIGISWLGIRSLRELSGLALIHNTHLCFVHTV 480  
421 DLSVFQNLQVIRGRILHNGAYSLTLOGIGISWLGIRSLRELSGLALIHNTHLCFVHTV 480  
Db 421 DLSVFQNLQVIRGRILHNGAYSLTLOGIGISWLGIRSLRELSGLALIHNTHLCFVHTV 480  
QY 481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540  
481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540  
Db 481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540  
QY 541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600  
541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600  
Db 541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600

601 PSQVPELSTYMPIMKFPDEBGAQCPCPINCTHSQVLDLDDKGCAPAEQASPLTISIISAVG 660  
601 PSQVPELSTYMPIMKFPDEBGAQCPCPINCTHSQVLDLDDKGCAPAEQASPLTISIISAVG 660  
QY 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRLIKETEL 720  
DB 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRLIKETEL 720  
QY 721 RKVAVLSSGAGCTYVKGIMIPDGENVKIPVAIKYLRNTSPKANKEIIDEAYVMAVGVS 780  
DB 721 RKVAVLSSGAGCTYVKGIMIPDGENVKIPVAIKYLRNTSPKANKEIIDEAYVMAVGVS 780  
QY 781 YVSRLLGICLTSTVQVLTQMLPYGCLDHYENRGRLSQDLNMQOIAKMSYLEDVR 840  
DB 781 YVSRLLGICLTSTVQVLTQMLPYGCLDHYENRGRLSQDLNMQOIAKMSYLEDVR 840  
QY 841 LVHRDLAARNVAVKSPNHVKITDFGLARLDDIDETRYHADGKVPDKMALESILRRFT 900  
DB 841 LVHRDLAARNVAVKSPNHVKITDFGLARLDDIDETRYHADGKVPDKMALESILRRFT 900  
QY 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGBRLPQPICTIDVYMIWVKCM 960  
DB 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGBRLPQPICTIDVYMIWVKCM 960  
QY 961 IDSECRPFRELVSFEFSMARDPQRFVVIQNEDLGPASPLDSTFYRSILLEDMDGLVDA 1020  
DB 961 IDSECRPFRELVSFEFSMARDPQRFVVIQNEDLGPASPLDSTFYRSILLEDMDGLVDA 1020  
QY 1021 BEYLVPOQGFPCPDPAFGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPSE 1080  
DB 1021 BEYLVPOQGFPCPDPAFGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPSE 1080  
QY 1081 AGSDVPDDDLGMAKGIQSLPTHPSPLOQYSEDPVPLBSETDGYAAPLTCSPQRPYV 1140  
DB 1081 AGSDVPDDDLGMAKGIQSLPTHPSPLOQYSEDPVPLBSETDGYAAPLTCSPQRPYV 1140  
QY 1141 NOPDVRPOPSPREBPLPAARPAKATLERPKLISGKGVKVDVAFGAGVNPPEYLTPO 1200  
DB 1141 NOPDVRPOPSPREBPLPAARPAKATLERPKLISGKGVKVDVAFGAGVNPPEYLTPO 1200  
QY 1201 GGAAPQHPHPAPFSAFDNLVYWDQPPERGAPSTFGKPTPAENPEYLGIDVPV 1255  
DB 1201 GGAAPQHPHPAPFSAFDNLVYWDQPPERGAPSTFGKPTPAENPEYLGIDVPV 1255

RESULT 7  
US-10-418-027-3  
; Sequence 3, Application US/10418027  
; Publication No. US20030224467A1  
; GENERAL INFORMATION:  
; APPLICANT: Osborne, C. Kent  
; APPLICANT: Schief, Rachel  
; APPLICANT: Bardou, Valerie  
; APPLICANT: Hilsenbeck, Susan  
; APPLICANT: Clark, Gary  
; APPLICANT: Wong, Jiemin  
; APPLICANT: Chamness, Gary  
; APPLICANT: Hopp, Torsten  
; TITLE OF INVENTION: AIB 1 as a prognostic marker and predictor of endocrine therapy  
; TITLE OF INVENTION: resistance  
; FILE REFERENCE: HO-P02396US1  
; CURRENT APPLICATION NUMBER: US/10/418,027  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: US 60/373,237  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 3  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Human  
US-10-418-027-3

Query Match 100.0%; Score 6815; DB 4; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MELAALRWGILLALLPPGAASSTVCTGTMDKRLPLSPSETHLMDLHLVGGCCVQVGNL 60  
DB 1 MELAALRWGILLALLPPGAASSTVCTGTMDKRLPLSPSETHLMDLHLVGGCCVQVGNL 60  
QY 61 ELTLPFNASLSPQDIQEVQGYVLIANQVQVPLRLRIVRGSTQFPEDNYALAVLDNG 120  
DB 61 ELTLPFNASLSPQDIQEVQGYVLIANQVQVPLRLRIVRGSTQFPEDNYALAVLDNG 120  
QY 121 DPLNNTTPVTGASDGLREIQLRSLEILKGVLIQBNPOLCYODTILMKDI FHKNNOLA 180  
DB 121 DPLNNTTPVTGASDGLREIQLRSLEILKGVLIQBNPOLCYODTILMKDI FHKNNOLA 180  
QY 181 LTLIDTRSRACHCSMCKSGSRGWSSSBDCOSLTFTVCAGGARCKGPLPTDCHEQC 240  
DB 181 LTLIDTRSRACHCSMCKSGSRGWSSSBDCOSLTFTVCAGGARCKGPLPTDCHEQC 240  
QY 241 AAGCTGPKHSDCLAFHNSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYAC 300  
DB 241 AAGCTGPKHSDCLAFHNSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYAC 300  
QY 301 YNYLSTDVSGCTIVCPHNOEVTAEDETQRCCKSKPCARVCYGLGMEHLREVAVTSAN 360  
DB 301 YNYLSTDVSGCTIVCPHNOEVTAEDETQRCCKSKPCARVCYGLGMEHLREVAVTSAN 360  
QY 361 IOEPAGCKKIFGSLAFPESEFDGPASNTAPLOEBOI QVPELTLEITGYLISAMP 420  
DB 361 IOEPAGCKKIFGSLAFPESEFDGPASNTAPLOEBOI QVPELTLEITGYLISAMP 420  
QY 421 DLSVFQNLQVIRGILHNGAVSLTLOGISLWLSRLSRLSGSLALIHNTHLCPVTV 480  
DB 421 DLSVFQNLQVIRGILHNGAVSLTLOGISLWLSRLSRLSGSLALIHNTHLCPVTV 480  
QY 481 PMDOLFNPQALHTANRDEECVGEGLACHOLCARGHCWGPPTQCVNCSQFLRGEC 540  
DB 481 PMDOLFNPQALHTANRDEECVGEGLACHOLCARGHCWGPPTQCVNCSQFLRGEC 540  
QY 541 VEECRVYQGLPREVYNAHCLPCHPBECOPONGSTTCGPREDDQCVACHYNDPFCVAC 600  
DB 541 VEECRVYQGLPREVYNAHCLPCHPBECOPONGSTTCGPREDDQCVACHYNDPFCVAC 600  
QY 601 PSQVPELSTYMPIMKFPDEBGAQCPCPINCTHSQVLDLDDKGCAPAEQASPLTISIISAVG 660  
DB 601 PSQVPELSTYMPIMKFPDEBGAQCPCPINCTHSQVLDLDDKGCAPAEQASPLTISIISAVG 660  
QY 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRLIKETEL 720  
DB 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRLIKETEL 720  
QY 721 RKVAVLSSGAGCTYVKGIMIPDGENVKIPVAIKYLRNTSPKANKEIIDEAYVMAVGVS 780  
DB 721 RKVAVLSSGAGCTYVKGIMIPDGENVKIPVAIKYLRNTSPKANKEIIDEAYVMAVGVS 780  
QY 781 YVSRLLGICLTSTVQVLTQMLPYGCLDHYENRGRLSQDLNMQOIAKMSYLEDVR 840  
DB 781 YVSRLLGICLTSTVQVLTQMLPYGCLDHYENRGRLSQDLNMQOIAKMSYLEDVR 840  
QY 841 LVHRDLAARNVAVKSPNHVKITDFGLARLDDIDETRYHADGKVPDKMALESILRRFT 900  
DB 841 LVHRDLAARNVAVKSPNHVKITDFGLARLDDIDETRYHADGKVPDKMALESILRRFT 900  
QY 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGBRLPQPICTIDVYMIWVKCM 960  
DB 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGBRLPQPICTIDVYMIWVKCM 960  
QY 961 IDSECRPFRELVSFEFSMARDPQRFVVIQNEDLGPASPLDSTFYRSILLEDMDGLVDA 1020  
DB 961 IDSECRPFRELVSFEFSMARDPQRFVVIQNEDLGPASPLDSTFYRSILLEDMDGLVDA 1020

QY 1021 BEYLVPQOGFCPCDDPAPAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPBEG 1080  
DB 1021 BEYLVPQOGFCPCDDPAPAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPBEG 1080  
QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLOQRYSEDPVPLPSETDGYAPLTCSPOPEVY 1140  
DB 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLOQRYSEDPVPLPSETDGYAPLTCSPOPEVY 1140  
QY 1141 NOPDVRPOPSPREBPLPAARPAAGATLERPKTLSPGKNGVVDVAFGAVENPEYLTPQ 1200  
DB 1141 NOPDVRPOPSPREBPLPAARPAAGATLERPKTLSPGKNGVVDVAFGAVENPEYLTPQ 1200  
QY 1201 GGAAPQHPHPPAFSPAFDNLVYWDODPPERGAPESTFKGTPTAENPEYLGLDVPV 1255  
DB 1201 GGAAPQHPHPPAFSPAFDNLVYWDODPPERGAPESTFKGTPTAENPEYLGLDVPV 1255  
  
RESULT 8  
US-10-394-322A-17  
; Sequence 17, Application US/10394322A  
; Publication No. US2003023291A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.  
; APPLICANT: Prescott, John C.  
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
; FILE REFERENCE: 39750-0006 US  
; CURRENT APPLICATION NUMBER: US/10/394,322A  
; CURRENT FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: US 60/366,892  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: FaestSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 1255  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-10-394-322A-17  
  
Query Match 100.0%; Score 6815; DB 4; Length 1255;  
Beet Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MELALACRWGLLALLPAGASTOVCTGDMKRLPASPETHLDMRLHYOCQVVOGNTL 60  
DB 1 MELALACRWGLLALLPAGASTOVCTGDMKRLPASPETHLDMRLHYOCQVVOGNTL 60  
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIANQVRQVPLQRLIVRGTOQLFEDNYALAVLDNG 120  
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLIANQVRQVPLQRLIVRGTOQLFEDNYALAVLDNG 120  
QY 121 DPLANTPTPTGASPGGLARLQRLSLTELKGGVLIQRPOLCYOPTIIMKQIFHNQOLA 180  
DB 121 DPLANTPTPTGASPGGLARLQRLSLTELKGGVLIQRPOLCYOPTIIMKQIFHNQOLA 180  
QY 121 DPLANTPTPTGASPGGLARLQRLSLTELKGGVLIQRPOLCYOPTIIMKQIFHNQOLA 180  
DB 121 DPLANTPTPTGASPGGLARLQRLSLTELKGGVLIQRPOLCYOPTIIMKQIFHNQOLA 180  
QY 181 LTLIDTNSBRACHPCSPMKGSRGWSSEBDCOSLTRTVACAGACARCKPLPTDCBECQ 240  
DB 181 LTLIDTNSBRACHPCSPMKGSRGWSSEBDCOSLTRTVACAGACARCKPLPTDCBECQ 240  
QY 241 AAGCTGPKXSDCLACLFHNSGICELHCPALVTYNTDTRESMPNPEGRTYFGASCVTACP 300  
DB 241 AAGCTGPKXSDCLACLFHNSGICELHCPALVTYNTDTRESMPNPEGRTYFGASCVTACP 300  
QY 301 YNYLSTDVGSCTLVCPRLHNOEYTAEDGTORCKRSPCARVCYGLGMEHLREVRVTSAN 360  
DB 301 YNYLSTDVGSCTLVCPRLHNOEYTAEDGTORCKRSPCARVCYGLGMEHLREVRVTSAN 360  
QY 361 IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLQPELOQVETLEBITGYLISAPDSLP 420  
DB 361 IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLQPELOQVETLEBITGYLISAPDSLP 420  
QY 421 DLSVFQNTQVIRGRILHNGAYSULTQGLGISWLGRLSRLSELGSGALIHNTHLCFVHTV 480  
DB 421 DLSVFQNTQVIRGRILHNGAYSULTQGLGISWLGRLSRLSELGSGALIHNTHLCFVHTV 480

QY 481 PMDQLFRNPQALHTANRPEDECEVGEGLACHQLCARGHCWGPGETQCVCNCSQFLRGEC 540  
DB 481 PMDQLFRNPQALHTANRPEDECEVGEGLACHQLCARGHCWGPGETQCVCNCSQFLRGEC 540  
QY 541 VEECEVLOGLPREYVNAHCLPCHEPCOPQNSVTCFGEADQCYACHYKDPPECVAC 600  
DB 541 VEECEVLOGLPREYVNAHCLPCHEPCOPQNSVTCFGEADQCYACHYKDPPECVAC 600  
QY 601 PEGYVPLDLSYMPIMKPEPBEAGCOPCPINCHTSCVDLDDKGPABORASPLTSISAVYG 660  
DB 601 PEGYVPLDLSYMPIMKPEPBEAGCOPCPINCHTSCVDLDDKGPABORASPLTSISAVYG 660  
QY 661 ILVVVLGVFEGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQOQMRILKETEL 720  
DB 661 ILVVVLGVFEGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQOQMRILKETEL 720  
QY 721 RKVYLVSAGAFGTVYKGIWIPDGENVKI PVAIKVLRNTSPKANKELIDEAYVMAGVSP 780  
DB 721 RKVYLVSAGAFGTVYKGIWIPDGENVKI PVAIKVLRNTSPKANKELIDEAYVMAGVSP 780  
QY 781 YVSRLLGICTSTVOLNOLMPYGLLDHVRNRRGLSQDILNMCQIAGMSYLEDVR 840  
DB 781 YVSRLLGICTSTVOLNOLMPYGLLDHVRNRRGLSQDILNMCQIAGMSYLEDVR 840  
QY 841 LVHRDLAARNVLYKSPNHVKTDPGLARLDDIDETRYADGKVPDKMMALESIARRFT 900  
DB 841 LVHRDLAARNVLYKSPNHVKTDPGLARLDDIDETRYADGKVPDKMMALESIARRFT 900  
QY 901 HQSDVMSYGVTVWELMTFGAKPYDGI PAREIPDLLEKBERLPQPICTIDVYMIWKCWM 960  
DB 901 HQSDVMSYGVTVWELMTFGAKPYDGI PAREIPDLLEKBERLPQPICTIDVYMIWKCWM 960  
QY 961 IDSECRPRFRELVSFSSMAADPORFVYIQNEDLGPAAPLSTFRSLLEDMDMDLVA 1020  
DB 961 IDSECRPRFRELVSFSSMAADPORFVYIQNEDLGPAAPLSTFRSLLEDMDMDLVA 1020  
QY 1021 BEYLVPQOGFCPCDDPAPAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPBEG 1080  
DB 1021 BEYLVPQOGFCPCDDPAPAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPBEG 1080  
QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLOQRYSEDPVPLPSETDGYAPLTCSPOPEVY 1140  
DB 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLOQRYSEDPVPLPSETDGYAPLTCSPOPEVY 1140  
QY 1141 NOPDVRPOPSPREBPLPAARPAAGATLERPKTLSPGKNGVVDVAFGAVENPEYLTPQ 1200  
DB 1141 NOPDVRPOPSPREBPLPAARPAAGATLERPKTLSPGKNGVVDVAFGAVENPEYLTPQ 1200  
QY 1201 GGAAPQHPHPPAFSPAFDNLVYWDODPPERGAPESTFKGTPTAENPEYLGLDVPV 1255  
DB 1201 GGAAPQHPHPPAFSPAFDNLVYWDODPPERGAPESTFKGTPTAENPEYLGLDVPV 1255  
  
RESULT 9  
US-10-245-871-553  
; Sequence 553, Application US/10245871  
; Publication No. US20030235594A1  
; GENERAL INFORMATION:  
; APPLICANT: HUMPHREYS, ROBERT  
; APPLICANT: XU, MINZHEN  
; TITLE OF INVENTION: 11-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES  
; FILE REFERENCE: RH-2013  
; CURRENT APPLICATION NUMBER: US/10/245,871  
; CURRENT FILING DATE: 2003-01-09  
; PRIOR APPLICATION NUMBER: 10/197,000  
; PRIOR FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: 09/396,813  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 905  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 553  
; LENGTH: 1255

TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-245-871-553

Query Match 100.0%; Score 6815; DB 4; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MELALACRMGLLALLPFGAASVCTGCTGDMKRLPASPEETHLDMRLHYOGCCVQVQNL 60
DB 1 MELALACRMGLLALLPFGAASVCTGCTGDMKRLPASPEETHLDMRLHYOGCCVQVQNL 60
QY 61 ELTYLPTNASLSFLDIOEVQGYVLIANQVQVPLQRLRYRGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLDIOEVQGYVLIANQVQVPLQRLRYRGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTSASPGGLRELQRLSTELIKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
DB 121 DPLNNTTPTVTSASPGGLRELQRLSTELIKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTYCAGGACRCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTYCAGGACRCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDTPESMPNDEGRYTGASCVTACP 300
DB 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDTPESMPNDEGRYTGASCVTACP 300
QY 301 YNYLSTDVSGCTLVCPPLHNOETVTAEDGTORCEKSKPCARVYGI GMEHLREVRVTSAN 360
DB 301 YNYLSTDVSGCTLVCPPLHNOETVTAEDGTORCEKSKPCARVYGI GMEHLREVRVTSAN 360
QY 361 IOEPAGCKKIFGSLAFLEPSPFDGASNTAPLQPEOLOFETLEETIGLYISAMPDILP 420
DB 361 IOEPAGCKKIFGSLAFLEPSPFDGASNTAPLQPEOLOFETLEETIGLYISAMPDILP 420
QY 421 DLVSFQNLQVIRGRILHNGAVSLTLQGLGISWLGRLSRELGSGLALIHNNTHLCFVATV 480
DB 421 DLVSFQNLQVIRGRILHNGAVSLTLQGLGISWLGRLSRELGSGLALIHNNTHLCFVATV 480
QY 481 PWDQLPKRNPHQALLHTANRPEDECCVGEGLAGHQLCARHGCMWPGTQCNCGQFLRGQEC 540
DB 481 PWDQLPKRNPHQALLHTANRPEDECCVGEGLAGHQLCARHGCMWPGTQCNCGQFLRGQEC 540
QY 541 VEECGVLQGLPREYVNAHCLPCHPECOPONGSVTCFGBAQCACAHYKDPFCVARG 600
DB 541 VEECGVLQGLPREYVNAHCLPCHPECOPONGSVTCFGBAQCACAHYKDPFCVARG 600
QY 601 PGGVPRDLSTYMPIMKFPDEEGACQCPINCTHSCVDLDDKGPAPQORASPLTSII SAVVG 660
DB 601 PGGVPRDLSTYMPIMKFPDEEGACQCPINCTHSCVDLDDKGPAPQORASPLTSII SAVVG 660
QY 661 ILLVVVLGVVFGILIKRROQKIRKTYMRRLLOETELVEPLTPSGAMPNOQMRILKETEL 720
DB 661 ILLVVVLGVVFGILIKRROQKIRKTYMRRLLOETELVEPLTPSGAMPNOQMRILKETEL 720
QY 721 RUVKVLGSGAGFTYVKGIMIPGENYKIPVAIKVLRENTSPKANKELIDEAVMAGVSP 780
DB 721 RUVKVLGSGAGFTYVKGIMIPGENYKIPVAIKVLRENTSPKANKELIDEAVMAGVSP 780
QY 781 YVSRLLGLICTSTVQLMPGYCLLDHVRNENRGLSGODLLNMCQIAKMSYLEDVR 840
DB 781 YVSRLLGLICTSTVQLMPGYCLLDHVRNENRGLSGODLLNMCQIAKMSYLEDVR 840
QY 841 LVHRDLAARNVLKSPBNHYKITDFFGLARLLDIDETRYHADGSKVPIKMMALESILRRFT 900
DB 841 LVHRDLAARNVLKSPBNHYKITDFFGLARLLDIDETRYHADGSKVPIKMMALESILRRFT 900
QY 901 HOSDWVSGVTVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVMIWVKCM 960
DB 901 HOSDWVSGVTVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVMIWVKCM 960
QY 961 IDSECRPRFRELVSERFMRADPQRFVVIQNEIDLGPASPLDSTFYRSLEDDMDGLVDA 1020
```

```
DB 961 IDSECRPRFRELVSERFMRADPQRFVVIQNEIDLGPASPLDSTFYRSLEDDMDGLVDA 1020
QY 1021 BEYLVPQGFPCPPAPAGAGMWHRRSSSTRSGGGDLTLGLEPSEEBAPRSLAPSE 1080
DB 1021 BEYLVPQGFPCPPAPAGAGMWHRRSSSTRSGGGDLTLGLEPSEEBAPRSLAPSE 1080
QY 1081 AGSDVFDGLGMGAAGLQSLPTHDPSPLOQRYSEDPVPLPSESTDGYVAPLTCSPQPEYV 1140
DB 1081 AGSDVFDGLGMGAAGLQSLPTHDPSPLOQRYSEDPVPLPSESTDGYVAPLTCSPQPEYV 1140
QY 1141 NOPDVRQPPSPRGRGPIPAARPAATLERPTLSFGKNGVYKVYAFAGAVNBEYLTPQ 1200
DB 1141 NOPDVRQPPSPRGRGPIPAARPAATLERPTLSFGKNGVYKVYAFAGAVNBEYLTPQ 1200
QY 1201 GGAAPQHPHPPAPFADNLYWMDQDPERGAPSTFGPTAENPEYLGIDVY 1255
DB 1201 GGAAPQHPHPPAPFADNLYWMDQDPERGAPSTFGPTAENPEYLGIDVY 1255
```

```
RESULT 10
US-10-469-162-3
; Sequence 3, Application US/10469162
; Publication No. US20040052811A1
; GENERAL INFORMATION:
; APPLICANT: Ziefelski, Christoph
; APPLICANT: Behamberger, Hubert
; APPLICANT: Breiteneder, Helmo
; APPLICANT: Jensen-Jarolim, Erika
; APPLICANT: Scheiner, Otto
; TITLE OF INVENTION: Vaccines Against Cancerous Diseases Associated With the HER-2/neu
; FILE REFERENCE: K 38 132/3yv
; CURRENT APPLICATION NUMBER: US/10/469,162
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: PCT/EP02/02111
; PRIOR FILING DATE: 2002-02-27
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(675)
; OTHER INFORMATION: Extracellular Domain
US-10-469-162-3
```

Query Match 100.0%; Score 6815; DB 4; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MELALACRMGLLALLPFGAASVCTGCTGDMKRLPASPEETHLDMRLHYOGCCVQVQNL 60
DB 1 MELALACRMGLLALLPFGAASVCTGCTGDMKRLPASPEETHLDMRLHYOGCCVQVQNL 60
QY 61 ELTYLPTNASLSFLDIOEVQGYVLIANQVQVPLQRLRYRGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLDIOEVQGYVLIANQVQVPLQRLRYRGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTSASPGGLRELQRLSTELIKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
DB 121 DPLNNTTPTVTSASPGGLRELQRLSTELIKGGVLIQRNPOLCYQDTIIMKDI FHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTYCAGGACRCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTYCAGGACRCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDTPESMPNDEGRYTGASCVTACP 300
```

Db 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNEGRTTFGASCVTACP 300  
Qy 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRATSAN 360  
Db 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRATSAN 360  
Qy 361 IOEPAGCKKIFGSLAFPLPESPDGDPASNTAPLOPEOLOVFETLEETITGLYISAMPDSL 420  
Db 361 IOEPAGCKKIFGSLAFPLPESPDGDPASNTAPLOPEOLOVFETLEETITGLYISAMPDSL 420  
Qy 421 DLSVFQNLQVIRGRILHNGAYSILTQGLGISMGLRSRLSGLALIHNTHLCFVHTV 480  
Db 421 DLSVFQNLQVIRGRILHNGAYSILTQGLGISMGLRSRLSGLALIHNTHLCFVHTV 480  
Qy 481 PWDOLFRRPHQALLHTANRPEDECVBEGLAGHQLCARHGWMPGPPTQCNCGQPLRGQEC 540  
Db 481 PWDOLFRRPHQALLHTANRPEDECVBEGLAGHQLCARHGWMPGPPTQCNCGQPLRGQEC 540  
Qy 541 VBECHVLQGLPREYVNAHCLPCHPECOPOGNSVTCFGEADQCVACHYKDPPECVANC 600  
Db 541 VBECHVLQGLPREYVNAHCLPCHPECOPOGNSVTCFGEADQCVACHYKDPPECVANC 600  
Qy 601 PSGVPRDLSYMPIMKFPDEBGAQCPCPINCTHSCVDLDDKGCPRABORASPLTSIIISAVG 660  
Db 601 PSGVPRDLSYMPIMKFPDEBGAQCPCPINCTHSCVDLDDKGCPRABORASPLTSIIISAVG 660  
Qy 661 ILLVVLGVVFGIILKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720  
Db 661 ILLVVLGVVFGIILKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720  
Qy 721 RKVKVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIELDEAYVMAGVSP 780  
Db 721 RKVKVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIELDEAYVMAGVSP 780  
Qy 781 YVSRLLGICLTSTVQLVTLQMPYGCILDHVRENRGLSQDILLNMQIYAKMSYLEDR 840  
Db 781 YVSRLLGICLTSTVQLVTLQMPYGCILDHVRENRGLSQDILLNMQIYAKMSYLEDR 840  
Qy 841 LVHRLLAARNVLVKSPNHVKITDPGLARLLDIDETRYHADGGKVPKXMALESILRRRT 900  
Db 841 LVHRLLAARNVLVKSPNHVKITDPGLARLLDIDETRYHADGGKVPKXMALESILRRRT 900  
Qy 901 HOSDWMSYGVTWELMTFGAKPYDGI PARBIPDLLEKGERLPOPPICTIDVTIMVCKWM 960  
Db 901 HOSDWMSYGVTWELMTFGAKPYDGI PARBIPDLLEKGERLPOPPICTIDVTIMVCKWM 960  
Qy 961 IDSECRPRERELVSERSMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDODMGDLVDA 1020  
Db 961 IDSECRPRERELVSERSMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDODMGDLVDA 1020  
Qy 1021 EBYLVPOQGFPCPDPAFGAGVHHRRSSSTRSGGDLITGLBSEBEAPRSPPLAPSRG 1080  
Db 1021 EBYLVPOQGFPCPDPAFGAGVHHRRSSSTRSGGDLITGLBSEBEAPRSPPLAPSRG 1080  
Qy 1081 AGSDVFDGDLGKAAGLQSLPTHPSPLOKRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
Db 1081 AGSDVFDGDLGKAAGLQSLPTHPSPLOKRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140  
Qy 1141 NOPDVRPOPSPREGLPAPAPAGATLERPKTILSPKINGVAVKAPFGAVENPEYLTPQ 1200  
Db 1141 NOPDVRPOPSPREGLPAPAPAGATLERPKTILSPKINGVAVKAPFGAVENPEYLTPQ 1200  
Qy 1201 GGAAQPHPPAPAFADNLTYWDOPPRRGAAPSTFKGTPTAENPEYLGDLVVP 1255  
Db 1201 GGAAQPHPPAPAFADNLTYWDOPPRRGAAPSTFKGTPTAENPEYLGDLVVP 1255

RESULT 11

US-10-253-286-553  
; Sequence 353, Application US/10253286  
; Publication No. US20040058881A1  
; GENERAL INFORMATION:  
; APPLICANT: HUMPHREYS, ROBERT

; APPLICANT: XU, MINZHEN  
; TITLE OF INVENTION: 11-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES  
; FILE REFERENCE: REH-2015  
; CURRENT APPLICATION NUMBER: US/10/253,286  
; PRIOR FILING DATE: 2003-01-13  
; PRIOR APPLICATION NUMBER: 10/197,000  
; PRIOR FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: 09/396,813  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 905  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 553  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-253-286-553

Query Match 100.0%; Score 6815; DB 4; Length 1255;

Best Local Similarity 100.0%; Pred. No; 0; Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 METALACRWGILLALLPRGAASSTOVCTGTDMLRLPASPEITLDMRLHLYGCGVVGVL 60  
Db 1 METALACRWGILLALLPRGAASSTOVCTGTDMLRLPASPEITLDMRLHLYGCGVVGVL 60  
Qy 61 ELTYLPTNASISFLQDIOEVQGYVLIAHQVQVPLQRLIVRGTLFEDNYALAVLDNG 120  
Db 61 ELTYLPTNASISFLQDIOEVQGYVLIAHQVQVPLQRLIVRGTLFEDNYALAVLDNG 120  
Qy 121 DPLNNTPTVYASPGGLRELOLRSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNQOLA 180  
Db 121 DPLNNTPTVYASPGGLRELOLRSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNQOLA 180  
Qy 181 LTLITNRSRACHPSSPMCKSGRCWGESSSDQSLTRTVACGACARCGPLPTDCHEQC 240  
Db 181 LTLITNRSRACHPSSPMCKSGRCWGESSSDQSLTRTVACGACARCGPLPTDCHEQC 240  
Qy 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNEGRTTFGASCVTACP 300  
Db 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNEGRTTFGASCVTACP 300  
Qy 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRATSAN 360  
Db 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRATSAN 360  
Qy 361 IOEPAGCKKIFGSLAFPLPESPDGDPASNTAPLOPEOLOVFETLEETITGLYISAMPDSL 420  
Db 361 IOEPAGCKKIFGSLAFPLPESPDGDPASNTAPLOPEOLOVFETLEETITGLYISAMPDSL 420  
Qy 421 DLSVFQNLQVIRGRILHNGAYSILTQGLGISMGLRSRLSGLALIHNTHLCFVHTV 480  
Db 421 DLSVFQNLQVIRGRILHNGAYSILTQGLGISMGLRSRLSGLALIHNTHLCFVHTV 480  
Qy 481 PWDOLFRRPHQALLHTANRPEDECVBEGLAGHQLCARHGWMPGPPTQCNCGQPLRGQEC 540  
Db 481 PWDOLFRRPHQALLHTANRPEDECVBEGLAGHQLCARHGWMPGPPTQCNCGQPLRGQEC 540  
Qy 541 VBECHVLQGLPREYVNAHCLPCHPECOPOGNSVTCFGEADQCVACHYKDPPECVANC 600  
Db 541 VBECHVLQGLPREYVNAHCLPCHPECOPOGNSVTCFGEADQCVACHYKDPPECVANC 600  
Qy 601 PSGVPRDLSYMPIMKFPDEBGAQCPCPINCTHSCVDLDDKGCPRABORASPLTSIIISAVG 660  
Db 601 PSGVPRDLSYMPIMKFPDEBGAQCPCPINCTHSCVDLDDKGCPRABORASPLTSIIISAVG 660  
Qy 661 ILLVVLGVVFGIILKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720  
Db 661 ILLVVLGVVFGIILKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720  
Qy 721 RKVKVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIELDEAYVMAGVSP 780  
Db 721 RKVKVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIELDEAYVMAGVSP 780

Qy 761 YVSRLLGICLSTVQVLTQMPYGCILDHVRENGRLGSDLLNMCMQIANGMSYLEBVR 840  
Db 761 YVSRLLGICLSTVQVLTQMPYGCILDHVRENGRLGSDLLNMCMQIANGMSYLEBVR 840  
Qy 841 LVHRDLAARNLVKSPNHNKTTDFGLARLLDIDETEVHADGKVIKMMALESLIRREF 900  
Db 841 LVHRDLAARNLVKSPNHNKTTDFGLARLLDIDETEVHADGKVIKMMALESLIRREF 900  
Qy 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPPOPICTIDVYMIWVKCM 960  
Db 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPPOPICTIDVYMIWVKCM 960  
Qy 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGPAAPLSTFYRSLLEDDMDGLVDA 1020  
Db 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGPAAPLSTFYRSLLEDDMDGLVDA 1020  
Qy 1021 EBYLVPQGGFCPPDAPAGAGVHHRHSSSTRSGGDLTGLPSEEBEAPRSLAPSEG 1080  
Db 1021 EBYLVPQGGFCPPDAPAGAGVHHRHSSSTRSGGDLTGLPSEEBEAPRSLAPSEG 1080  
Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSDPTVPLPSETDGYVAPLTCSFPQRYV 1140  
Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSDPTVPLPSETDGYVAPLTCSFPQRYV 1140  
Qy 1141 NOPDVRPOPSPRSGPLPAARPAATLERPKTLPFGKGVVDVAFGGAIVENPEYLTPO 1200  
Db 1141 NOPDVRPOPSPRSGPLPAARPAATLERPKTLPFGKGVVDVAFGGAIVENPEYLTPO 1200  
Qy 1201 GGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAAPSTFGTPTAENPEYLGIDVPV 1255  
Db 1201 GGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAAPSTFGTPTAENPEYLGIDVPV 1255

## RESULT 12

US-10-762-128-6  
; Sequence 6, Application US/10762128  
; Publication No. US20040219161A1  
; GENERAL INFORMATION:  
; APPLICANT: Scholler, Nathalie B.  
; APPLICANT: Dieis, Mary L.  
; APPLICANT: Hellstrom, Inggerd  
; APPLICANT: Hellstrom, Karl Erik  
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES  
; FILE REFERENCE: 730033.409C1  
; CURRENT APPLICATION NUMBER: US/10/762,128  
; PRIOR FILING DATE: 2004-01-20  
; PRIOR APPLICATION NUMBER: US 09/441,411  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-762-128-6

Query Match 100.0%; Score 6815; DB 5; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MELAALCWMGLLALLPFGAASVCTGTDMKLRIPASBETHLDMLRHLYGSCQVYQCNL 60  
Db 1 MELAALCWMGLLALLPFGAASVCTGTDMKLRIPASBETHLDMLRHLYGSCQVYQCNL 60  
Qy 61 ELTYLPTNASLSFLODIOEVQGVYLIANQVQVPLQGLRYRGQLPEDNYALAVLNG 120  
Db 61 ELTYLPTNASLSFLODIOEVQGVYLIANQVQVPLQGLRYRGQLPEDNYALAVLNG 120  
Qy 121 DPLNNTTVYTGASPGGLRELQRLSTELIKGGVLIQRNPOLCYOPTIIMKQIFHKQNOA 180  
Db 121 DPLNNTTVYTGASPGGLRELQRLSTELIKGGVLIQRNPOLCYOPTIIMKQIFHKQNOA 180  
Qy 181 LTLIDTNRSRACHPCSPCKGSRCWGESSEDCQSLITRTVCAGGCARCKGPLPTDCHEQC 240

Db 181 LTLIDTNRSRACHPCSPCKGSRCWGESSEDCQSLITRTVCAGGCARCKGPLPTDCHEQC 240  
Qy 241 AAGCTGPHSDCLACLPNHSIGICELHCPALVTYNTPTFFSMPNBEGRYTGASCVTAC 300  
Db 241 AAGCTGPHSDCLACLPNHSIGICELHCPALVTYNTPTFFSMPNBEGRYTGASCVTAC 300  
Qy 301 YNVLSTVGSCTLVCPHNOEVTABDGTORCEKSKCARVCYGLGMEHLREVAVTSAN 360  
Db 301 YNVLSTVGSCTLVCPHNOEVTABDGTORCEKSKCARVCYGLGMEHLREVAVTSAN 360  
Qy 361 IOEFAKCKITFGSLAPLPESFDGPASNTAPLOEBOLOVEFTLEBITGYLISAMPB 420  
Db 361 IOEFAKCKITFGSLAPLPESFDGPASNTAPLOEBOLOVEFTLEBITGYLISAMPB 420  
Qy 421 DLSVFONLQVIRGILHNGAYSLTLOGLISMLGRSLRELSGSLALIHNTHLCEFTV 480  
Db 421 DLSVFONLQVIRGILHNGAYSLTLOGLISMLGRSLRELSGSLALIHNTHLCEFTV 480  
Qy 481 PMDQLFRNPQALHTANRPEDECVGGLACHQLCARGHCWGPPTQCVCNSQFLRGQC 540  
Db 481 PMDQLFRNPQALHTANRPEDECVGGLACHQLCARGHCWGPPTQCVCNSQFLRGQC 540  
Qy 541 VEECRVITQGLPREVYNAHCLPCHPECOPONGSYTCGPRADCCVACHYDPPFCVARC 600  
Db 541 VEECRVITQGLPREVYNAHCLPCHPECOPONGSYTCGPRADCCVACHYDPPFCVARC 600  
Qy 601 PSGVKPDLSTYMPYKFPDEEGACOPCPINCHSCVDDLDKCPAEORASPLTISIAYVG 660  
Db 601 PSGVKPDLSTYMPYKFPDEEGACOPCPINCHSCVDDLDKCPAEORASPLTISIAYVG 660  
Qy 661 ILVVLGVVFGILIKRQOKIRKYMRLLQETLEVEPLTPSGAMPNQAQRIKTEL 720  
Db 661 ILVVLGVVFGILIKRQOKIRKYMRLLQETLEVEPLTPSGAMPNQAQRIKTEL 720  
Qy 721 RKVYVLSGAGTYTKGIVIPDGENVKIPVAIKYLRNTSPKAKETILDEAYVMAAGVSP 780  
Db 721 RKVYVLSGAGTYTKGIVIPDGENVKIPVAIKYLRNTSPKAKETILDEAYVMAAGVSP 780  
Qy 781 YVSRLLGICLSTVQVLTQMPYGCILDHVRENGRLGSDLLNMCMQIANGMSYLEBVR 840  
Db 781 YVSRLLGICLSTVQVLTQMPYGCILDHVRENGRLGSDLLNMCMQIANGMSYLEBVR 840  
Qy 841 LVHRDLAARNLVKSPNHNKTTDFGLARLLDIDETEVHADGKVIKMMALESLIRREF 900  
Db 841 LVHRDLAARNLVKSPNHNKTTDFGLARLLDIDETEVHADGKVIKMMALESLIRREF 900  
Qy 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPPOPICTIDVYMIWVKCM 960  
Db 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPPOPICTIDVYMIWVKCM 960  
Qy 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGPAAPLSTFYRSLLEDDMDGLVDA 1020  
Db 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGPAAPLSTFYRSLLEDDMDGLVDA 1020  
Qy 1021 EBYLVPQGGFCPPDAPAGAGVHHRHSSSTRSGGDLTGLPSEEBEAPRSLAPSEG 1080  
Db 1021 EBYLVPQGGFCPPDAPAGAGVHHRHSSSTRSGGDLTGLPSEEBEAPRSLAPSEG 1080  
Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSDPTVPLPSETDGYVAPLTCSFPQRYV 1140  
Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSDPTVPLPSETDGYVAPLTCSFPQRYV 1140  
Qy 1141 NOPDVRPOPSPRSGPLPAARPAATLERPKTLPFGKGVVDVAFGGAIVENPEYLTPO 1200  
Db 1141 NOPDVRPOPSPRSGPLPAARPAATLERPKTLPFGKGVVDVAFGGAIVENPEYLTPO 1200  
Qy 1201 GGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAAPSTFGTPTAENPEYLGIDVPV 1255  
Db 1201 GGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAAPSTFGTPTAENPEYLGIDVPV 1255

## RESULT 13

```

US-10-484-067-1
; Sequence 1, Application US/10484067
; Publication No. US2004024166A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF CALIFORNIA
; APPLICANT: NELSON, Edward L.
; TITLE OF INVENTION: HER2/NEU TARGET ANTIGEN AND USE OF SAME TO STIMULATE AN IMMUNE RE
; FILE REFERENCE: UCI1170-1
; CURRENT APPLICATION NUMBER: US/10/484,067
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: PCT/US02/22975
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/306,250
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-484-067-1

Query Match      100.0%; Score 6815; DB 5; Length 1255;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1      MELALACRWGLLALLPPGAASVQVCTGTDMLRLPASPETHLDMRLHYOGCCVVGNTL 60
DB      1      MELALACRWGLLALLPPGAASVQVCTGTDMLRLPASPETHLDMRLHYOGCCVVGNTL 60
QY      61      ELYLPTNASISFLDIOEVOGYVLIANOVQVPLQRLIRYRGTOLEFEDNYALAVLNG 120
DB      61      ELYLPTNASISFLDIOEVOGYVLIANOVQVPLQRLIRYRGTOLEFEDNYALAVLNG 120
QY      121      DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPQLCYODTIIMKDI FHKNOQLA 180
DB      121      DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPQLCYODTIIMKDI FHKNOQLA 180
QY      181      LTLIDTNSRACHPSSPMCKSGRCSGSESDCSLTRVYCGAGCARCKGRLPTDCCHQC 240
DB      181      LTLIDTNSRACHPSSPMCKSGRCSGSESDCSLTRVYCGAGCARCKGRLPTDCCHQC 240
QY      241      AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFEPMNPEGRTYFGASCVTACP 300
DB      241      AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFEPMNPEGRTYFGASCVTACP 300
QY      301      YNYLSTDVGSCTLVCPPLHNOEVTAEQTCRCEKSKPCARVCYGLGMEHLREVRATVSAN 360
DB      301      YNYLSTDVGSCTLVCPPLHNOEVTAEQTCRCEKSKPCARVCYGLGMEHLREVRATVSAN 360
QY      361      IOEPAGCKKIFGSLAFPLPESFGDPAASNTAPLOPEQLOVFETLEBITGLYLSAMPDSL 420
DB      361      IOEPAGCKKIFGSLAFPLPESFGDPAASNTAPLOPEQLOVFETLEBITGLYLSAMPDSL 420
QY      421      DLSVQNTQVIRGRILHNGAYSLTQGLISWLGRLRELSGSLALHNNTHLCFVHTV 480
DB      421      DLSVQNTQVIRGRILHNGAYSLTQGLISWLGRLRELSGSLALHNNTHLCFVHTV 480
QY      481      FMDQLFRNPHQALHTANRPEDECVGEGLACHQLCARGHGMGPPTOCVNCGSQLRGQEC 540
DB      481      FMDQLFRNPHQALHTANRPEDECVGEGLACHQLCARGHGMGPPTOCVNCGSQLRGQEC 540
QY      541      VEECVLQGLPREYVNAHCLFCHPECQPNQSVTCFGEADQCYACAHYKDPPECVAVRC 600
DB      541      VEECVLQGLPREYVNAHCLFCHPECQPNQSVTCFGEADQCYACAHYKDPPECVAVRC 600
QY      601      PSGVPRDLSYMPIMKFPDEBGACOPCINCTHSCVDLDDKGPABORASPLTSII SAYVG 660
DB      601      PSGVPRDLSYMPIMKFPDEBGACOPCINCTHSCVDLDDKGPABORASPLTSII SAYVG 660
QY      661      ILLVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEBLTPSGAMPNOQMRILKETEL 720
DB      661      ILLVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEBLTPSGAMPNOQMRILKETEL 720

```

```

QY      721      RKVKVLGSGAFGTVYKGIWIPDGENVKIPVALKVLRENTSPKANKIILDEAYVMAGVGP 780
DB      721      RKVKVLGSGAFGTVYKGIWIPDGENVKIPVALKVLRENTSPKANKIILDEAYVMAGVGP 780
QY      781      YVSRLLIGLITSTVQLTQMLPBYGCLLDHVRNRRGLSQDILLANCMQIAKMSYLEDR 840
DB      781      YVSRLLIGLITSTVQLTQMLPBYGCLLDHVRNRRGLSQDILLANCMQIAKMSYLEDR 840
QY      841      LVHRDLAARNVLYKSPNNVKTIDREGALLDIDETRYHADGKVPILKMALESILRRRT 900
DB      841      LVHRDLAARNVLYKSPNNVKTIDREGALLDIDETRYHADGKVPILKMALESILRRRT 900
QY      901      HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRLEPPPICTIDVYIMVCKM 960
DB      901      HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRLEPPPICTIDVYIMVCKM 960
QY      961      IDSECRPRRELVSFSSRNARDPQRFVYIQNEDLGPASPILDSTFYRSLLEDMDGLVDA 1020
DB      961      IDSECRPRRELVSFSSRNARDPQRFVYIQNEDLGPASPILDSTFYRSLLEDMDGLVDA 1020
QY      1021      ERYLVPQGFPCPDPAFGAGVHHRSSSTRSGGDLTLGLESEEBEAPRSLAPSRG 1080
DB      1021      ERYLVPQGFPCPDPAFGAGVHHRSSSTRSGGDLTLGLESEEBEAPRSLAPSRG 1080
QY      1081      AGSDVFDDGLGKAAGKLSLPTHDPSPLQRYSEDPTVPLPSETGYVAPLTCSPQPEYV 1140
DB      1081      AGSDVFDDGLGKAAGKLSLPTHDPSPLQRYSEDPTVPLPSETGYVAPLTCSPQPEYV 1140
QY      1141      NOPDVRPOPSPRREGPLPAARPAATLERPKTLSPGKGVVVDVAFGAVENPEYLTPO 1200
DB      1141      NOPDVRPOPSPRREGPLPAARPAATLERPKTLSPGKGVVVDVAFGAVENPEYLTPO 1200
QY      1201      GGAAPQHPPPAFSPAFNLYWDDPPRKGAPSTFKGTPTAENPEYLGDLVPV 1255
DB      1201      GGAAPQHPPPAFSPAFNLYWDDPPRKGAPSTFKGTPTAENPEYLGDLVPV 1255

RESULT 14
US-10-723-860-9
; Sequence 9, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natsaba
; APPLICANT: Ginsburg, Wendy M.
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882, 0193, NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patent version 3.2
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-9

Query Match      100.0%; Score 6815; DB 5; Length 1255;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1      MELALACRWGLLALLPPGAASVQVCTGTDMLRLPASPETHLDMRLHYOGCCVVGNTL 60
DB      1      MELALACRWGLLALLPPGAASVQVCTGTDMLRLPASPETHLDMRLHYOGCCVVGNTL 60
QY      61      ELYLPTNASISFLDIOEVOGYVLIANOVQVPLQRLIRYRGTOLEFEDNYALAVLNG 120
DB      61      ELYLPTNASISFLDIOEVOGYVLIANOVQVPLQRLIRYRGTOLEFEDNYALAVLNG 120

```



QY 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTILMKDI FHKNNOLA 180  
DB 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTILMKDI FHKNNOLA 180  
QY 181 LTLIDTNSRACHPCSPMKSGSRCSGSESSDCOSLTRVCGAGGRCGRLPTDCCHQC 240  
DB 181 LTLIDTNSRACHPCSPMKSGSRCSGSESSDCOSLTRVCGAGGRCGRLPTDCCHQC 240  
QY 241 AAGCTGPGHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNBEGRYTFGASCVTACP 300  
DB 241 AAGCTGPGHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNBEGRYTFGASCVTACP 300  
QY 301 YNYLSTDVGSCTLVCPILHNOEVTABDGTQRCCKSKPCARVCYGLGMHLREVRVTSAN 360  
DB 301 YNYLSTDVGSCTLVCPILHNOEVTABDGTQRCCKSKPCARVCYGLGMHLREVRVTSAN 360  
QY 361 IOEFAGCKKITFGSLAFLPESFPGDASNTAPLOPEQLOVFETLEBITGYLYISAMPDSL 420  
DB 361 IOEFAGCKKITFGSLAFLPESFPGDASNTAPLOPEQLOVFETLEBITGYLYISAMPDSL 420  
QY 421 DLSVFQNLQVIRGRIIHNGAYSILTQGLISWLGRLSRLSGSLALIHNTHLCFVHTV 480  
DB 421 DLSVFQNLQVIRGRIIHNGAYSILTQGLISWLGRLSRLSGSLALIHNTHLCFVHTV 480  
QY 481 PWDOLFRRNPQALHTTANRPEDECVGEGIACHOLCARGHCMGPGPTQCVCNCSQFLRGQC 540  
DB 481 PWDOLFRRNPQALHTTANRPEDECVGEGIACHOLCARGHCMGPGPTQCVCNCSQFLRGQC 540  
QY 541 VEEBCVLOGLPREYVNAHCLPCHPECOPONGSVTCFGBEADQVCAAHYKDPFCVARC 600  
DB 541 VEEBCVLOGLPREYVNAHCLPCHPECOPONGSVTCFGBEADQVCAAHYKDPFCVARC 600  
QY 601 PSGVPRDLSYMPIMKFPDEBGACQPCPINCTHSCVDLDDKGPABQASPLTISIISAVVG 660  
DB 601 PSGVPRDLSYMPIMKFPDEBGACQPCPINCTHSCVDLDDKGPABQASPLTISIISAVVG 660  
QY 661 ILLVVIVGVFGIILIKRRQOKIRKTYMRLLQETELVEBLTPSGAMPQOMRIKELTEL 720  
DB 661 ILLVVIVGVFGIILIKRRQOKIRKTYMRLLQETELVEBLTPSGAMPQOMRIKELTEL 720  
QY 721 RKVKVLGSGAGRTYKGIWIPDGENVKI PVAIKVLRENTSPKANKEIIDEAIVMAGVSP 780  
DB 721 RKVKVLGSGAGRTYKGIWIPDGENVKI PVAIKVLRENTSPKANKEIIDEAIVMAGVSP 780  
QY 781 YVSRLLGICLSTVOLVTQLMFYGCLLDHVRNRRGLSGODLLNMCQI IAKMSLYEDVR 840  
DB 781 YVSRLLGICLSTVOLVTQLMFYGCLLDHVRNRRGLSGODLLNMCQI IAKMSLYEDVR 840  
QY 841 LVHRDLAARNVL VKSPNFKITD FGLARLLDIDETETHADGGKVP I KMALESILRRFT 900  
DB 841 LVHRDLAARNVL VKSPNFKITD FGLARLLDIDETETHADGGKVP I KMALESILRRFT 900  
QY 901 HOSDVWSYGVYTWELMTFGAKPYDGI PAKEI PDLLEKGERLPQPICTIDVMIVMCKM 960  
DB 901 HOSDVWSYGVYTWELMTFGAKPYDGI PAKEI PDLLEKGERLPQPICTIDVMIVMCKM 960  
QY 961 IDSECRPFRELIVSEFSRMAPDQRFVYIQNEDLGPASPLDSTFFRSLIEDDDMDLVDYA 1020  
DB 961 IDSECRPFRELIVSEFSRMAPDQRFVYIQNEDLGPASPLDSTFFRSLIEDDDMDLVDYA 1020  
QY 1021 EBYLVPOQGFPCPDAPGAGVWHRHRSSTRSGGDLTLGLEPSEBEAPRSLAPSBG 1080  
DB 1021 EBYLVPOQGFPCPDAPGAGVWHRHRSSTRSGGDLTLGLEPSEBEAPRSLAPSBG 1080  
QY 1081 AGSDVFPDDDLGMAKGLQSLPTHDPSPLQRYSEBPTVPLPSETGUYAVALTCSQPEYV 1140  
DB 1081 AGSDVFPDDDLGMAKGLQSLPTHDPSPLQRYSEBPTVPLPSETGUYAVALTCSQPEYV 1140  
QY 1141 NOPDVRPOPSPBREGPLPAARPAAGATLEBRKTLSPGKNGVNDVFAFGAVENPEYLTPQ 1200  
DB 1141 NOPDVRPOPSPBREGPLPAARPAAGATLEBRKTLSPGKNGVNDVFAFGAVENPEYLTPQ 1200  
QY 1201 GGAAPQPHPPAPFSPAFDNLVYWDODPPERGAPPESTFKGTPAENPEYLGDLDPV 1255

DB 1201 GGAAPQPHPPAPFSPAFDNLVYWDODPPERGAPPESTFKGTPAENPEYLGDLDPV 1255  
RESULT 15  
US-10-871-708-9  
; Sequence 9, Application US/10871708  
; Publication No. US2005018186A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiang, Chih-Sheng  
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED  
; TITLE OF INVENTION: ANTIGENS IN COMPOSITIONS FOR VARIOUS TYPES OF CANCERS  
; FILE REFERENCE: MANAK-035A  
; CURRENT APPLICATION NUMBER: US/10/871,708  
; PRIOR FILING DATE: 2004-06-17  
; PRIOR FILING DATE: 2003-06-17  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Receptor protein-tyrosine kinase  
US-10-871-708-9  
Query Match 100.0%; Score 6815; DB 5; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MELAALCRWGLLLALLPRGASTQVCTGTMKRLRLPASPTHLDMLRLHYOGCOVVOGNTL 60  
DB 1 MELAALCRWGLLLALLPRGASTQVCTGTMKRLRLPASPTHLDMLRLHYOGCOVVOGNTL 60  
QY 61 ELTYLPTNASLSFLQDIOEVGGVLIANNOVROVPLQRLIVRGTLQFEDNYALAVLDNG 120  
DB 61 ELTYLPTNASLSFLQDIOEVGGVLIANNOVROVPLQRLIVRGTLQFEDNYALAVLDNG 120  
QY 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTILMKDI FHKNNOLA 180  
DB 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTILMKDI FHKNNOLA 180  
QY 181 LTLIDTNSRACHPCSPMKSGSRCSGSESSDCOSLTRVCGAGGRCGRLPTDCCHQC 240  
DB 181 LTLIDTNSRACHPCSPMKSGSRCSGSESSDCOSLTRVCGAGGRCGRLPTDCCHQC 240  
QY 241 AAGCTGPGHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNBEGRYTFGASCVTACP 300  
DB 241 AAGCTGPGHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNBEGRYTFGASCVTACP 300  
QY 301 YNYLSTDVGSCTLVCPILHNOEVTABDGTQRCCKSKPCARVCYGLGMHLREVRVTSAN 360  
DB 301 YNYLSTDVGSCTLVCPILHNOEVTABDGTQRCCKSKPCARVCYGLGMHLREVRVTSAN 360  
QY 361 IOEFAGCKKITFGSLAFLPESFPGDASNTAPLOPEQLOVFETLEBITGYLYISAMPDSL 420  
DB 361 IOEFAGCKKITFGSLAFLPESFPGDASNTAPLOPEQLOVFETLEBITGYLYISAMPDSL 420  
QY 421 DLSVFQNLQVIRGRIIHNGAYSILTQGLISWLGRLSRLSGSLALIHNTHLCFVHTV 480  
DB 421 DLSVFQNLQVIRGRIIHNGAYSILTQGLISWLGRLSRLSGSLALIHNTHLCFVHTV 480  
QY 481 PWDOLFRRNPQALHTTANRPEDECVGEGIACHOLCARGHCMGPGPTQCVCNCSQFLRGQC 540  
DB 481 PWDOLFRRNPQALHTTANRPEDECVGEGIACHOLCARGHCMGPGPTQCVCNCSQFLRGQC 540  
QY 541 VEEBCVLOGLPREYVNAHCLPCHPECOPONGSVTCFGBEADQVCAAHYKDPFCVARC 600  
DB 541 VEEBCVLOGLPREYVNAHCLPCHPECOPONGSVTCFGBEADQVCAAHYKDPFCVARC 600  
QY 601 PSGVPRDLSYMPIMKFPDEBGACQPCPINCTHSCVDLDDKGPABQASPLTISIISAVVG 660  
DB 601 PSGVPRDLSYMPIMKFPDEBGACQPCPINCTHSCVDLDDKGPABQASPLTISIISAVVG 660





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2006, 11:08:37 ; Search time 50 Seconds  
(without alignments)  
2075.160 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 METALCRWGLLALLPPGA.....TFKGTPTANPEYLGIDVIV 1255

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database : Issued Patents AA:

1: /cgn2\_6/prodata/1/1aa/5 COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/6 COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/7 COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/8 COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/9 COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/BACKFILE1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6815	100.0	1255	1 US-08-625-101-2	Sequence 2, Appl 1
2	6815	100.0	1255	1 US-08-356-786-2	Sequence 2, Appl 1
3	6815	100.0	1255	2 US-09-441-411-6	Sequence 6, Appl 1
4	6815	100.0	1255	2 US-09-167-516-2	Sequence 2, Appl 1
5	6806	99.9	1255	1 US-08-467-083-68	Sequence 68, Appl 1
6	6806	99.9	1255	1 US-08-414-417B-68	Sequence 68, Appl 1
7	6806	99.9	1255	1 US-08-486-348A-68	Sequence 68, Appl 1
8	6806	99.9	1255	1 US-08-468-545B-68	Sequence 68, Appl 1
9	6806	99.9	1255	2 US-08-466-680B-68	Sequence 68, Appl 1
10	6806	99.9	1255	2 US-09-527-487-2	Sequence 2, Appl 1
11	6806	99.9	1255	2 US-09-811-115-3	Sequence 3, Appl 1
12	6806	99.9	1255	2 US-09-354-533-68	Sequence 8, Appl 1
13	6732	98.8	1255	1 US-08-484-438-8	Sequence 4, Appl 1
14	3630	53.3	782	1 US-09-146-283-4	Sequence 4, Appl 1
15	3630	53.3	782	1 US-08-579-823A-4	Sequence 4, Appl 1
16	3630	53.3	782	2 US-09-344-195-4	Sequence 4, Appl 1
17	3590	52.7	645	2 US-09-602-812A-13	Sequence 13, Appl 1
18	3473	51.0	624	2 US-08-422-108-1	Sequence 1, Appl 1
19	3473	51.0	624	2 US-08-422-734-1	Sequence 1, Appl 1
20	3169	46.5	1210	1 US-08-475-035-4	Sequence 7, Appl 1
21	3167	46.5	1210	2 US-08-484-438-7	Sequence 7, Appl 1
22	3167	46.5	1210	2 US-09-715-249-2	Sequence 7, Appl 1
23	3146	46.2	1210	2 US-09-723-307-67	Sequence 67, Appl 1
24	3091	45.4	580	1 US-08-414-417B-69	Sequence 69, Appl 1
25	3091	45.4	580	1 US-08-486-348A-69	Sequence 69, Appl 1
26	3091	45.4	580	1 US-08-468-545B-69	Sequence 69, Appl 1
27	3091	45.4	580	2 US-08-466-680B-69	Sequence 69, Appl 1

28	3091	45.4	580	2 US-09-354-533-69	Sequence 69, Appl 1
29	3004.5	44.1	1308	1 US-08-484-438-2	Sequence 2, Appl 1
30	2871.5	42.1	1058	1 US-08-484-438-4	Sequence 4, Appl 1
31	2441.5	35.8	1342	1 US-07-978-895-4	Sequence 4, Appl 1
32	2441.5	35.8	1342	1 US-08-484-438-9	Sequence 9, Appl 1
33	2441.5	35.8	1342	1 US-08-473-119-4	Sequence 4, Appl 1
34	2441.5	35.8	1342	1 US-08-475-352-4	Sequence 4, Appl 1
35	2441.5	35.8	1342	2 US-09-170-699-4	Sequence 4, Appl 1
36	2441.5	35.8	1360	2 US-09-949-016-8022	Sequence 8022, Appl 1
37	2434	35.7	1343	6 5183884-4	Patent No. 5183884
38	1878	27.6	419	2 US-09-630-155-2	Sequence 2, Appl 1
39	1646.5	24.2	705	1 US-08-456-647B-4	Sequence 4, Appl 1
40	1646.5	24.2	705	1 US-08-237-401A-4	Sequence 4, Appl 1
41	1532.5	22.5	644	1 US-08-336-708A-9	Sequence 9, Appl 1
42	1451	21.3	911	1 US-08-484-438-10	Sequence 10, Appl 1
43	1389.5	20.4	265	1 US-07-857-224B-66	Sequence 66, Appl 1
44	1235	18.1	279	1 US-08-701-191A-14	Sequence 14, Appl 1
45	1235	18.1	279	2 US-09-664-526-14	Sequence 14, Appl 1

## ALIGNMENTS

RESULT 1  
US-08-625-101-2  
; Sequence 2, Application US/08625101  
; Patent No. 5869445  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE  
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION  
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PC-DOS/MS-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/625,101  
; FILING DATE: 01-APR-1996  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C7  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-625-101-2

Query Match 100.0%; Score 6815; DB 1; Length 1255;  
Best Local Similarity 100.0%; Pred. No: 0;  
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALCRWGLLALLPPGAASVQCTGTDMKRLRLPASPEFTLDMRLHYOCGCVQGNL 60  
DB 1 METALCRWGLLALLPPGAASVQCTGTDMKRLRLPASPEFTLDMRLHYOCGCVQGNL 60

```

QY 61 ELYTPTNASISFLDIOEVQGYVLIANQVROVPLQRLIRVGTOLFEDNYALAVLNG 120
DB 61 ELYTPTNASISFLDIOEVQGYVLIANQVROVPLQRLIRVGTOLFEDNYALAVLNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSITELIKGGVLIQRNPOLCYQDTTIIMKDIFFHKNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSITELIKGGVLIQRNPOLCYQDTTIIMKDIFFHKNOLA 180
QY 181 LTLIDTNSRACHPSPMCKSGRCWGESSEDCQSITRTVACAGGARGCKGPLTDCCHBCC 240
DB 181 LTLIDTNSRACHPSPMCKSGRCWGESSEDCQSITRTVACAGGARGCKGPLTDCCHBCC 240
QY 241 AAGCTGPKHSDCLAFHNSGICBLHCPALVTYNTDTFESMPNEGRYTFGASCVTIACP 300
DB 241 AAGCTGPKHSDCLAFHNSGICBLHCPALVTYNTDTFESMPNEGRYTFGASCVTIACP 300
QY 301 YNYLSTDVQSCITLVCPLHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
DB 301 YNYLSTDVQSCITLVCPLHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
QY 361 IOEPAGCKKIFGSLAFLPESFDPDASNTAPLOPELOVFEETLEETGYLSAMPDULP 420
DB 361 IOEPAGCKKIFGSLAFLPESFDPDASNTAPLOPELOVFEETLEETGYLSAMPDULP 420
QY 421 DLSVFQNLQVIRGRILHNGAVSLTLOGLISWLGRLSRLGSGALIHNTHLCFVHTV 480
DB 421 DLSVFQNLQVIRGRILHNGAVSLTLOGLISWLGRLSRLGSGALIHNTHLCFVHTV 480
QY 481 PWDOLFRRPHQALLHTANRPEDECVBEGGLACHQOLARHGCWPGTQVCNCSQFLRGQEC 540
DB 481 PWDOLFRRPHQALLHTANRPEDECVBEGGLACHQOLARHGCWPGTQVCNCSQFLRGQEC 540
QY 541 VEECEVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCAVCAHYKPPCVCARC 600
DB 541 VEECEVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCAVCAHYKPPCVCARC 600
QY 601 PSGVAPDLSYMBIMKFPDEBGAQOPCINCTHSCVDLDDKGPAPORASPLTISIAYVG 660
DB 601 PSGVAPDLSYMBIMKFPDEBGAQOPCINCTHSCVDLDDKGPAPORASPLTISIAYVG 660
QY 661 ILVVVVLGVVFGILLKRRQOKIRKTYTMRLLQETELVEPLTSGAMPNOQVRIKETEL 720
DB 661 ILVVVVLGVVFGILLKRRQOKIRKTYTMRLLQETELVEPLTSGAMPNOQVRIKETEL 720
QY 721 RRVKVLGSGAGFTVKKGIWIPDGENVKIPVAIKVIRENTSPANKELIDEAVYMAVGQSP 780
DB 721 RRVKVLGSGAGFTVKKGIWIPDGENVKIPVAIKVIRENTSPANKELIDEAVYMAVGQSP 780
QY 781 YVSRLLGICLTISTVOLVQLMPYGLLDHVRNRRGLSGODLLNMCQIAKMSYLEDVR 840
DB 781 YVSRLLGICLTISTVOLVQLMPYGLLDHVRNRRGLSGODLLNMCQIAKMSYLEDVR 840
QY 841 LVHRDLAARNVLVKSPPNHVKTIDFGLARLLDIDETRYHADGKVPITKMALESILRRPFT 900
DB 841 LVHRDLAARNVLVKSPPNHVKTIDFGLARLLDIDETRYHADGKVPITKMALESILRRPFT 900
QY 901 HOSDVWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLQOPPICTIDVYMIWVKCM 960
DB 901 HOSDVWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLQOPPICTIDVYMIWVKCM 960
QY 961 IDSECRPRRELIVSEFSRWARDPQRFVVIQNEEDLGASPLDSTFYRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRRELIVSEFSRWARDPQRFVVIQNEEDLGASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 EBYLVPOQGFCEPDPAFGAGVHHRHSSSTRSGGDLTGLLEPSEEBEAPRSPAPBSG 1080
DB 1021 EBYLVPOQGFCEPDPAFGAGVHHRHSSSTRSGGDLTGLLEPSEEBEAPRSPAPBSG 1080
QY 1081 AGSDVFDGDLGMAKAGLOSLPTHDPSPLORYSSEDPVPLPSETGCVYVPLTCSPOPEV 1140
DB 1081 AGSDVFDGDLGMAKAGLOSLPTHDPSPLORYSSEDPVPLPSETGCVYVPLTCSPOPEV 1140

```

```

QY 1141 NOPVRFQPPSPREGPLPAAPAGATLERPKTSLSPKNGVYKDVAFAGAVENPEYLTPO 1200
DB 1141 NOPVRFQPPSPREGPLPAAPAGATLERPKTSLSPKNGVYKDVAFAGAVENPEYLTPO 1200
QY 1201 GGAAPQHPAPFAPFADNLVYWDODPERGAPSTFGPTTAENPEYLGIDVPV 1255
DB 1201 GGAAPQHPAPFAPFADNLVYWDODPERGAPSTFGPTTAENPEYLGIDVPV 1255

RESULT 2
US-08-356-786-2
; Sequence 2, Application US/08356786
; Patent No. 587305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pletcher, Testra, Hurwitz, & Thibeault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; SOFTWARE: Patent in release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pletcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; TELEPHONE/DOCKET NUMBER: CRP-053
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-786-2

Query Match 100.0%; Score 6815; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELALCGWGLLALLPPGASSTOVCTGDMKRLPASPETHLDMRLHLYOGCOVVGSL 60
DB 1 MELALCGWGLLALLPPGASSTOVCTGDMKRLPASPETHLDMRLHLYOGCOVVGSL 60
QY 61 ELYTPTNASISFLDIOEVQGYVLIANQVROVPLQRLIRVGTOLFEDNYALAVLNG 120
DB 61 ELYTPTNASISFLDIOEVQGYVLIANQVROVPLQRLIRVGTOLFEDNYALAVLNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSITELIKGGVLIQRNPOLCYQDTTIIMKDIFFHKNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSITELIKGGVLIQRNPOLCYQDTTIIMKDIFFHKNOLA 180
QY 181 LTLIDTNSRACHPSPMCKSGRCWGESSEDCQSITRTVACAGGARGCKGPLTDCCHBCC 240
DB 181 LTLIDTNSRACHPSPMCKSGRCWGESSEDCQSITRTVACAGGARGCKGPLTDCCHBCC 240

```

```

Db 181 LTLIDTNRSRACHPSPCKSGSRGWSESSDCQSLTRTVACAGCARCKGPLETDCCHQC 240
Qy 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVTYNTDPFESMPNBEGRTPASCVTACP 300
Db 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVTYNTDPFESMPNBEGRTPASCVTACP 300
Qy 301 YNYLSTDVGSCTLVPLHNOEVTABEDGTORCEKSKPCARVCYGLGMEHLREVRVAVTSAN 360
Db 301 YNYLSTDVGSCTLVPLHNOEVTABEDGTORCEKSKPCARVCYGLGMEHLREVRVAVTSAN 360
Qy 361 IOEFAGCKKIFGSLAFLEPSFDGDPASNTAPLQPEOLQVFETLEITGYLISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFLEPSFDGDPASNTAPLQPEOLQVFETLEITGYLISAMPDSL 420
Qy 421 DLSVFQNTQVTRGRILHNGAVSLTLOGLGISWLGRLSRLBELSGSLALHHNTHLCFVHTV 480
Db 421 DLSVFQNTQVTRGRILHNGAVSLTLOGLGISWLGRLSRLBELSGSLALHHNTHLCFVHTV 480
Qy 481 PMDOLFRNPHOALHTTANRPEDECVGEGLAGHQLCARGHCMGPGFTQCVCNSQFLRGQEC 540
Db 481 PMDOLFRNPHOALHTTANRPEDECVGEGLAGHQLCARGHCMGPGFTQCVCNSQFLRGQEC 540
Qy 541 VEBECVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVANC 600
Db 541 VEBECVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVANC 600
Qy 601 PSQVPRDLSYMPIMKFPDEBGACQCPINCHSCVDLDKGCFAQORASPLTSIIISAVNG 660
Db 601 PSQVPRDLSYMPIMKFPDEBGACQCPINCHSCVDLDKGCFAQORASPLTSIIISAVNG 660
Qy 661 ILVVVVLGVFGIILKRRQOKIRKTYTMRLLQETELVEPLTPSGAMPNOAQRILIKETEL 720
Db 661 ILVVVVLGVFGIILKRRQOKIRKTYTMRLLQETELVEPLTPSGAMPNOAQRILIKETEL 720
Qy 721 RKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRBNTSPRANKIIDEAVYMAVGSP 780
Db 721 RKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRBNTSPRANKIIDEAVYMAVGSP 780
Qy 781 YVSRLLGICLTSTVOLVQLMYGCILHDYRENRGLSQDILLNMCQIAKMSYLEDR 840
Db 781 YVSRLLGICLTSTVOLVQLMYGCILHDYRENRGLSQDILLNMCQIAKMSYLEDR 840
Qy 841 LVHRLAARNVLYKSPNHVKITDPGLARLDDIDETEYHADGGKVPKMMALSSIIRRT 900
Db 841 LVHRLAARNVLYKSPNHVKITDPGLARLDDIDETEYHADGGKVPKMMALSSIIRRT 900
Qy 901 HOSDVSXGVTWELMTFGAKPYDGI PAEIPDLLEKGERLPQPICTIDVYMIWVKCM 960
Db 901 HOSDVSXGVTWELMTFGAKPYDGI PAEIPDLLEKGERLPQPICTIDVYMIWVKCM 960
Qy 961 IDSECRPRFRELVSFESRMAADPQRFVYQNEDLGASPLDSTFYRSLLEDMDGLDVA 1020
Db 961 IDSECRPRFRELVSFESRMAADPQRFVYQNEDLGASPLDSTFYRSLLEDMDGLDVA 1020
Qy 1021 BEVLYPQOGFCPPDAPAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSPPLAPSBG 1080
Db 1021 BEVLYPQOGFCPPDAPAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSPPLAPSBG 1080
Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140
Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEV 1140
Qy 1141 NOPDVRPOPSPRREGPLPAARPAAGATLEPKTLSPEKNGVVDVAFGAVENPEVLTQ 1200
Db 1141 NOPDVRPOPSPRREGPLPAARPAAGATLEPKTLSPEKNGVVDVAFGAVENPEVLTQ 1200
Qy 1201 GGAAPQHPHPAPSPADNLYYMDODPPERGAPESTFKGPTAENPEYLGLOVPV 1255
Db 1201 GGAAPQHPHPAPSPADNLYYMDODPPERGAPESTFKGPTAENPEYLGLOVPV 1255

```

RESULT 3  
US-09-441-411-6

```

; Sequence 6, Application US/09441411
; Patent No. 6734172
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary U.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FaSTSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-411-6

Query Match      100.0%; Score 6815; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No: 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MELAALCRWGLLALLPFGAASVQVCTGTDKRLRPASBETHLDMRLHYGCGVQGNL 60
Db 1 MELAALCRWGLLALLPFGAASVQVCTGTDKRLRPASBETHLDMRLHYGCGVQGNL 60
Qy 61 ELTYPTNASLFLDIDIOEVGYVLIANNQVQVLOLRITRGVQLPEDNALAVLDNG 120
Db 61 ELTYPTNASLFLDIDIOEVGYVLIANNQVQVLOLRITRGVQLPEDNALAVLDNG 120
Qy 121 DPLNNTPTVYASPGGLRELOLRSLTEILKGSVLIOBNPOLCYOPTILMKDIFHKNQOLA 180
Db 121 DPLNNTPTVYASPGGLRELOLRSLTEILKGSVLIOBNPOLCYOPTILMKDIFHKNQOLA 180
Qy 181 LTLIDTNRSRACHPSPCKSGSRGWSESSDCQSLTRTVACAGCARCKGPLETDCCHQC 240
Db 181 LTLIDTNRSRACHPSPCKSGSRGWSESSDCQSLTRTVACAGCARCKGPLETDCCHQC 240
Qy 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVTYNTDPFESMPNBEGRTPASCVTACP 300
Db 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVTYNTDPFESMPNBEGRTPASCVTACP 300
Qy 301 YNYLSTDVGSCTLVPLHNOEVTABEDGTORCEKSKPCARVCYGLGMEHLREVRVAVTSAN 360
Db 301 YNYLSTDVGSCTLVPLHNOEVTABEDGTORCEKSKPCARVCYGLGMEHLREVRVAVTSAN 360
Qy 361 IOEFAGCKKIFGSLAFLEPSFDGDPASNTAPLQPEOLQVFETLEITGYLISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFLEPSFDGDPASNTAPLQPEOLQVFETLEITGYLISAMPDSL 420
Qy 421 DLSVFQNTQVTRGRILHNGAVSLTLOGLGISWLGRLSRLBELSGSLALHHNTHLCFVHTV 480
Db 421 DLSVFQNTQVTRGRILHNGAVSLTLOGLGISWLGRLSRLBELSGSLALHHNTHLCFVHTV 480
Qy 481 PMDOLFRNPHOALHTTANRPEDECVGEGLAGHQLCARGHCMGPGFTQCVCNSQFLRGQEC 540
Db 481 PMDOLFRNPHOALHTTANRPEDECVGEGLAGHQLCARGHCMGPGFTQCVCNSQFLRGQEC 540
Qy 541 VEBECVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVANC 600
Db 541 VEBECVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVANC 600
Qy 601 PSQVPRDLSYMPIMKFPDEBGACQCPINCHSCVDLDKGCFAQORASPLTSIIISAVNG 660
Db 601 PSQVPRDLSYMPIMKFPDEBGACQCPINCHSCVDLDKGCFAQORASPLTSIIISAVNG 660
Qy 661 ILVVVVLGVFGIILKRRQOKIRKTYTMRLLQETELVEPLTPSGAMPNOAQRILIKETEL 720
Db 661 ILVVVVLGVFGIILKRRQOKIRKTYTMRLLQETELVEPLTPSGAMPNOAQRILIKETEL 720
Qy 721 RKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRBNTSPRANKIIDEAVYMAVGSP 780

```

Db 721 RKVVLGSGAFGTYYKGIWIPDGENVKI PVAIKVIRENTSPKANKELIDEAVYVAGVSP 780  
 QY 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHVENRGRSLGODLLNMCQIAGMSYLEBYR 840  
 Db 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHVENRGRSLGODLLNMCQIAGMSYLEBYR 840  
 QY 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEHADGKVPDKMMALLESILRRFT 900  
 Db 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEHADGKVPDKMMALLESILRRFT 900  
 QY 901 HQSDVWSGYVTWELMTFGAKPYDGI PAREIPDLLEKGERLPQPPICITIDYVMIMVKCM 960  
 Db 901 HQSDVWSGYVTWELMTFGAKPYDGI PAREIPDLLEKGERLPQPPICITIDYVMIMVKCM 960  
 QY 961 IDSECRPFRELVSFESRMARDPQRFVYI QNEDLPASPLDSTFYRSLEDDDMGLVDA 1020  
 Db 961 IDSECRPFRELVSFESRMARDPQRFVYI QNEDLPASPLDSTFYRSLEDDDMGLVDA 1020  
 QY 1021 EBYLVPQGGFCPPDAPAGAGVHHRRSSSTRSGGDLTLGLPSESEAPRSLAPSEG 1080  
 Db 1021 EBYLVPQGGFCPPDAPAGAGVHHRRSSSTRSGGDLTLGLPSESEAPRSLAPSEG 1080  
 QY 1081 AGSDVPDDDLGMAKAGIQLSPTHPSPLOQYSBDPTVPLBSETGVYVAPLTCSPQPEYV 1140  
 Db 1081 AGSDVPDDDLGMAKAGIQLSPTHPSPLOQYSBDPTVPLBSETGVYVAPLTCSPQPEYV 1140  
 QY 1141 NOPDVRPOPSPREBPLPAAPAGATLERPKTLSPGKGVVQDVAFGAGAVENPEYLTPO 1200  
 Db 1141 NOPDVRPOPSPREBPLPAAPAGATLERPKTLSPGKGVVQDVAFGAGAVENPEYLTPO 1200  
 QY 1201 GGAAPQHPPPAFSPAFNLYWDODPPERGAPSTFGKGTPTAENPEYLGIDVPY 1255  
 Db 1201 GGAAPQHPPPAFSPAFNLYWDODPPERGAPSTFGKGTPTAENPEYLGIDVPY 1255

# RESULT 4

US-09-167-516-2  
 Sequence 2, Application US/09167516  
 Patent No. 6953573

## GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.  
 APPLICANT: Disla, Mary L.  
 TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE  
 REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION  
 TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/167,516  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/625,101  
 FILING DATE: 01-APR-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629  
 REFERENCE/DOCKET NUMBER: 920010.448C7  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1255 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-167-516-2

Query Match 100.0%; Score 6815; DB 2; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MELAALCRWGLLLALLPPGAASVQCTGTDMKRLPASPETHLMDLRLYGGCCVQGNL 60  
 1 MELAALCRWGLLLALLPPGAASVQCTGTDMKRLPASPETHLMDLRLYGGCCVQGNL 60  
 61 ELTYLPNMASTFLODIQEVGVYVLIANOVROYPLORLIRVSGTOLFEDNYALAVLDNG 120  
 61 ELTYLPNMASTFLODIQEVGVYVLIANOVROYPLORLIRVSGTOLFEDNYALAVLDNG 120  
 121 DPLNNTTPVTGASPGGLREIQLRLSTELKGVLIQRNPOLCYODTILMKDI FHKNNQLA 180  
 121 DPLNNTTPVTGASPGGLREIQLRLSTELKGVLIQRNPOLCYODTILMKDI FHKNNQLA 180  
 181 LTLIDYTRSRACHPQSPCKGSCRCWGBSSDDCSLTRTVCAGGCAKCGPLPTDCHEQC 240  
 181 LTLIDYTRSRACHPQSPCKGSCRCWGBSSDDCSLTRTVCAGGCAKCGPLPTDCHEQC 240  
 241 AAGCTGPGHSDCLALCPHNSGICELHCPALVTYNTDPFESMPNBEGRYTGASCVTACP 300  
 241 AAGCTGPGHSDCLALCPHNSGICELHCPALVTYNTDPFESMPNBEGRYTGASCVTACP 300  
 301 YNVLSTDVGSCTLYVCPHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360  
 301 YNVLSTDVGSCTLYVCPHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360  
 361 IQEPAKCKKIFGSLAFIPESFDGPASNTAPLOEOLQVFTLEITGYLYISMPDGLP 420  
 361 IQEPAKCKKIFGSLAFIPESFDGPASNTAPLOEOLQVFTLEITGYLYISMPDGLP 420  
 421 DLSVFQNLQVIRGRILHNGAYSLTLOGIGTBMGLRSRLREGSGLALIHNTTHLCFYTV 480  
 421 DLSVFQNLQVIRGRILHNGAYSLTLOGIGTBMGLRSRLREGSGLALIHNTTHLCFYTV 480  
 481 PMDQLFNPHQALHTNRPDEDCVGGIACHQICAGHCWGPPTQCVNCSQFLRGEC 540  
 481 PMDQLFNPHQALHTNRPDEDCVGGIACHQICAGHCWGPPTQCVNCSQFLRGEC 540  
 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCGPEADQCVACHYDPPFCVARC 600  
 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCGPEADQCVACHYDPPFCVARC 600  
 601 PSQVAPDLSTYMPKPFDEBACOPCPINCTHSCVDLDDKCPAPORASPLTSIIISAVG 660  
 601 PSQVAPDLSTYMPKPFDEBACOPCPINCTHSCVDLDDKCPAPORASPLTSIIISAVG 660  
 661 ILVVVVGAVGILIKRQOKIRKTYMKRLLOETELVEPLTPSGAMPQAOAMRLKETEL 720  
 661 ILVVVVGAVGILIKRQOKIRKTYMKRLLOETELVEPLTPSGAMPQAOAMRLKETEL 720  
 721 RKVVLGSGAFGTYYKGIWIPDGENVKI PVAIKVIRENTSPKANKELIDEAVYVAGVSP 780  
 721 RKVVLGSGAFGTYYKGIWIPDGENVKI PVAIKVIRENTSPKANKELIDEAVYVAGVSP 780  
 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHVENRGRSLGODLLNMCQIAGMSYLEBYR 840  
 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHVENRGRSLGODLLNMCQIAGMSYLEBYR 840  
 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEHADGKVPDKMMALLESILRRFT 900  
 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEHADGKVPDKMMALLESILRRFT 900  
 901 HQSDVWSGYVTWELMTFGAKPYDGI PAREIPDLLEKGERLPQPPICITIDYVMIMVKCM 960

Db 901 HQSDWMSGVTVWELMTFGAKPYDGI PARBI PDLLEKGERLPQPPICITIDVYMIWKCMM 960  
 Qy 961 IDSECRPRFRELVSFESRMAADPORFVVIQNEIDLGPASPLDSTFYRSLLEDDMGDLVDA 1020  
 Db 961 IDSECRPRFRELVSFESRMAADPORFVVIQNEIDLGPASPLDSTFYRSLLEDDMGDLVDA 1020  
 Qy 1021 BEYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGGLTGLPSEBEARSPPLAPSEB 1080  
 Db 1021 BEYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGGLTGLPSEBEARSPPLAPSEB 1080  
 Qy 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPOPEYV 1140  
 Db 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPOPEYV 1140  
 Qy 1141 NQPDVRFPPSPBREGPLPAAPAGATLTERPPTLSFGKGVYKDVAFAGAVENPEYLTPO 1200  
 Db 1141 NQPDVRFPPSPBREGPLPAAPAGATLTERPPTLSFGKGVYKDVAFAGAVENPEYLTPO 1200  
 Qy 1201 GGAAPQHPAPSPAPFNDLYWDDPPERGAPESTFGTPTAENPEYLGDLVPV 1255  
 Db 1201 GGAAPQHPAPSPAPFNDLYWDDPPERGAPESTFGTPTAENPEYLGDLVPV 1255

## RESULT 5

US-08-467-083-68  
 Sequence 68, Application US/08467083  
 Patent No. 5726023  
 GENERAL INFORMATION:  
 APPLICANT: Cheever, Martin A.  
 APPLICANT: Disla, Mary L.  
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN  
 TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
 NUMBER OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seed and Berry  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: US  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/467,083  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/414,417  
 FILING DATE: 06-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629  
 REFERENCE/DOCKET NUMBER: 920010.448C2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 TELEX: 3723836 SEEDANBERY  
 INFORMATION FOR SEQ ID NO: 68:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1255 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-08-467-083-68

Query Match 99.9%; Score 6806; DB 1; Length 1255;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELAAACRMGLLALLPFGAASVQCTGCTGKRLRLPASPEITHLMDLRLYQSGCVQGNL 60  
 Db 1 MELAAACRMGLLALLPFGAASVQCTGCTGKRLRLPASPEITHLMDLRLYQSGCVQGNL 60  
 Qy 61 ELTYLPTNASLSFLODIOEVQGYVLIANQVROVPLQRLIRVGTOLFEDNYAALVDNG 120  
 Db 61 ELTYLPTNASLSFLODIOEVQGYVLIANQVROVPLQRLIRVGTOLFEDNYAALVDNG 120  
 Qy 121 DPLNNTTPVTGASGGGLREIQLRLSTELKGGVLIQRRPOLCYODPTLIMKDI FHKNNOLA 180  
 Db 121 DPLNNTTPVTGASGGGLREIQLRLSTELKGGVLIQRRPOLCYODPTLIMKDI FHKNNOLA 180  
 Qy 181 LTLIDTNRSRACHSPKCKSRCKWSSSEDCSLTRTVACGAGCARCKGPLPTDCHEQC 240  
 Db 181 LTLIDTNRSRACHSPKCKSRCKWSSSEDCSLTRTVACGAGCARCKGPLPTDCHEQC 240  
 Qy 241 AAGCTGPHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNBEGRYTTGASCVTACP 300  
 Db 241 AAGCTGPHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNBEGRYTTGASCVTACP 300  
 Qy 301 YNYLSTVGSCTVLCPLHNOEVTAEADGTQCEKCSKPCARVCYGLGMEHLREVAVTSAN 360  
 Db 301 YNYLSTVGSCTVLCPLHNOEVTAEADGTQCEKCSKPCARVCYGLGMEHLREVAVTSAN 360  
 Qy 361 IOEPAGCKIFGSLAFLPESFEDGPASNTAPLOEOLQVFETLEITGYLYISAMPDLP 420  
 Db 361 IOEPAGCKIFGSLAFLPESFEDGPASNTAPLOEOLQVFETLEITGYLYISAMPDLP 420  
 Qy 421 DLSVFQNLQVIRGILNNGAVSLTLOGLISWLGIRSLREISGLAIHNHTHLCFHTV 480  
 Db 421 DLSVFQNLQVIRGILNNGAVSLTLOGLISWLGIRSLREISGLAIHNHTHLCFHTV 480  
 Qy 481 PMDQLFRRPHQALHTNRPDECEVGEGLACHQICARHCMPGPTQCVCNCSOPIRGEC 540  
 Db 481 PMDQLFRRPHQALHTNRPDECEVGEGLACHQICARHCMPGPTQCVCNCSOPIRGEC 540  
 Qy 541 VEECRVLOGLPREVYNARHCPCHECOPQNGSVTCFGEPAQCVACAHYDPPCVARC 600  
 Db 541 VEECRVLOGLPREVYNARHCPCHECOPQNGSVTCFGEPAQCVACAHYDPPCVARC 600  
 Qy 601 PSQVVPDLSTYMPWKPFDEBGACQPCPINCTHSCVDLDDKCPABORASPLTISAVG 660  
 Db 601 PSQVVPDLSTYMPWKPFDEBGACQPCPINCTHSCVDLDDKCPABORASPLTISAVG 660  
 Qy 661 ILLVTVLGAVFGIILKRRQKIRKCTMRRLQETELVEPLTPSGAMPQAOAMRLIKETEL 720  
 Db 661 ILLVTVLGAVFGIILKRRQKIRKCTMRRLQETELVEPLTPSGAMPQAOAMRLIKETEL 720  
 Qy 721 RKVAVLGSAGFYVYKGIWIPDGENVKIPVAIKYLRENTSFKANKEILDEAVYVAGVSP 780  
 Db 721 RKVAVLGSAGFYVYKGIWIPDGENVKIPVAIKYLRENTSFKANKEILDEAVYVAGVSP 780  
 Qy 781 YVSRLLGICLSTVQVLTQMPYGCCLLDHVENRGLSGODLNMCMQIAKMSYLEBVR 840  
 Db 781 YVSRLLGICLSTVQVLTQMPYGCCLLDHVENRGLSGODLNMCMQIAKMSYLEBVR 840  
 Qy 841 LVHRDLAARNLVYKSPHNVKTTDGLARLLDIDETRYHADGKVPKIMMALESILRRFT 900  
 Db 841 LVHRDLAARNLVYKSPHNVKTTDGLARLLDIDETRYHADGKVPKIMMALESILRRFT 900  
 Qy 901 HQSDWMSGVTVWELMTFGAKPYDGI PARBI PDLLEKGERLPQPPICITIDVYMIWKCMM 960  
 Db 901 HQSDWMSGVTVWELMTFGAKPYDGI PARBI PDLLEKGERLPQPPICITIDVYMIWKCMM 960  
 Qy 961 IDSECRPRFRELVSFESRMAADPORFVVIQNEIDLGPASPLDSTFYRSLLEDDMGDLVDA 1020  
 Db 961 IDSECRPRFRELVSFESRMAADPORFVVIQNEIDLGPASPLDSTFYRSLLEDDMGDLVDA 1020  
 Qy 1021 BEYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGGLTGLPSEBEARSPPLAPSEB 1080  
 Db 1021 BEYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGGLTGLPSEBEARSPPLAPSEB 1080  
 Qy 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPOPEYV 1140



Db 1081 AGSVFPGDLMGAAKGLQSLPTHDPSLQRYSEDPVLPSETDGYVAPLTCSPQBEYV 1140  
 Qy 1141 NOPVRRPQPSRRGGLPAARPAAGATLERPKTLSPGKGVVVDVAFAGGAIVENBEYLTPO 1200  
 Db 1141 NOPVRRPQPSRRGGLPAARPAAGATLERPKTLSPGKGVVVDVAFAGGAIVENBEYLTPO 1200  
 Qy 1201 GGAAPQHPPPAPFPAFNDLYWMDQDPPERGAPSTFKGPTAENPEYLGIDVYV 1255  
 Db 1201 GGAAPQHPPPAPFPAFNDLYWMDQDPPERGAPSTFKGPTAENPEYLGIDVYV 1255

RESULT 6  
 US-08-414-117B-68  
 ; Sequence 68, Application US/08414417B  
 ; Patent No. 5801005  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Disla, Mary L.  
 ; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
 ; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
 ; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
 ; NUMBER OF SEQUENCES: 69  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Seed and Berry LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: US  
 ; ZIP: 98104-7092  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/414,417B  
 ; FILING DATE: 31-MAR-1995  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Sharkey, Richard G.  
 ; REGISTRATION NUMBER: 32,629  
 ; REFERENCE/DOCKET NUMBER: 920010.448C2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 622-4900  
 ; TELEFAX: (206) 682-6031  
 ; INFORMATION FOR SEQ ID NO: 68:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1255 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-414-117B-68

Query Match 99.9%; Score 6806; DB 1; Length 1255;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELAALCRWGLLLALPPGAASQVCTGTDKRLRPAPEETHLMDLRHLVGGCOVVGNTL 60  
 Db 1 MELAALCRWGLLLALPPGAASQVCTGTDKRLRPAPEETHLMDLRHLVGGCOVVGNTL 60  
 Qy 61 ELTYLPTNASISPLQDIOEVGVYLIANQVROYVLORLATVRGTQLFEDNYALAVLDNG 120  
 Db 61 ELTYLPTNASISPLQDIOEVGVYLIANQVROYVLORLATVRGTQLFEDNYALAVLDNG 120  
 Qy 121 DPLANNTTPVVGASGGGLRELQRLSLTELKGVLIQRRPOLCYQDITLIMKQIFRKNQOLA 180  
 Db 121 DPLANNTTPVVGASGGGLRELQRLSLTELKGVLIQRRPOLCYQDITLIMKQIFRKNQOLA 180  
 Qy 181 LTLIDTNRSRACHCSPKCKSGRCMGSESDQSLTRTVACGAGCARCKGSLPTDCCHQC 240  
 Db 181 LTLIDTNRSRACHCSPKCKSGRCMGSESDQSLTRTVACGAGCARCKGSLPTDCCHQC 240

Qy 241 AAGCTGKHSDDCLACHFNHSGICEHLCPALVTYNTDTPESMNPBESRYTFGASCVTACP 300  
 Db 241 AAGCTGKHSDDCLACHFNHSGICEHLCPALVTYNTDTPESMNPBESRYTFGASCVTACP 300  
 Qy 301 YNYISTDVGSCTIYVCPILHNOEVTAEQDTCRCKSKCAVVCVGLGMEHLREVAVLVSAN 360  
 Db 301 YNYISTDVGSCTIYVCPILHNOEVTAEQDTCRCKSKCAVVCVGLGMEHLREVAVLVSAN 360  
 Qy 361 IOEFAGCKKTFGSLAFPLESFPDGPASNTAPLPEQLQVEETELEEITGYLYISAMPDLP 420  
 Db 361 IOEFAGCKKTFGSLAFPLESFPDGPASNTAPLPEQLQVEETELEEITGYLYISAMPDLP 420  
 Qy 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLGISWGLSRSLRGLSLALIHNTHLCPVHTV 480  
 Db 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLGISWGLSRSLRGLSLALIHNTHLCPVHTV 480  
 Qy 481 PMDQLFENPHQALHTNANREDECVSGGLACHQICANGHCMGPPTOCVNCOSQPLRGQEC 540  
 Db 481 PMDQLFENPHQALHTNANREDECVSGGLACHQICANGHCMGPPTOCVNCOSQPLRGQEC 540  
 Qy 541 VEEGRVLOGLPREYVNAHGLPCHPECOPONGSVTCFGEPAQOCVACAHYKDPFCVARC 600  
 Db 541 VEEGRVLOGLPREYVNAHGLPCHPECOPONGSVTCFGEPAQOCVACAHYKDPFCVARC 600  
 Qy 601 PSGVKPDLSTMPYWKPFDEGACQPCPINCTHSCVDLDDYGCAPABOPASPLTISAVVG 660  
 Db 601 PSGVKPDLSTMPYWKPFDEGACQPCPINCTHSCVDLDDYGCAPABOPASPLTISAVVG 660  
 Qy 661 ILVYVTVGVVFGILIKRROOKIRKYMRLLOETELVEPLTPSGAMNOAMRLKTEL 720  
 Db 661 ILVYVTVGVVFGILIKRROOKIRKYMRLLOETELVEPLTPSGAMNOAMRLKTEL 720  
 Qy 721 RKVKVLSGAFGVYKGIWIPDENVKI PVAIKVLRNTPSPKANKELIDEAYVAVAGVSP 780  
 Db 721 RKVKVLSGAFGVYKGIWIPDENVKI PVAIKVLRNTPSPKANKELIDEAYVAVAGVSP 780  
 Qy 781 YVSRRLGICLTSTYQVLTQMPYGCILDHVRENKRGSGODLNMCMQIAKMSYLEYDVR 840  
 Db 781 YVSRRLGICLTSTYQVLTQMPYGCILDHVRENKRGSGODLNMCMQIAKMSYLEYDVR 840  
 Qy 841 LVHRDLAARVNLVKSPPNHVKTIDPGLARLLDIDETEVHAAGGVKPIKMALESTLRRFT 900  
 Db 841 LVHRDLAARVNLVKSPPNHVKTIDPGLARLLDIDETEVHAAGGVKPIKMALESTLRRFT 900  
 Qy 901 HQSDVMSYGVTVLWELMTFGAKPYDGI PAREIPDLLEKGERLPORPCTIDVYMIWVCM 960  
 Db 901 HQSDVMSYGVTVLWELMTFGAKPYDGI PAREIPDLLEKGERLPORPCTIDVYMIWVCM 960  
 Qy 961 IDSECRPRFRELVSERSMARDPQRFVYIQNEDLGPAASPLDSTFYRSLLEDDDMGDLVDA 1020  
 Db 961 IDSECRPRFRELVSERSMARDPQRFVYIQNEDLGPAASPLDSTFYRSLLEDDDMGDLVDA 1020  
 Qy 1021 EEYLVVPOQGFPCDPPAPGAGVMVHHRSSSTRSGGGLTLGLPEPSEBA PRSPLAPSEG 1080  
 Db 1021 EEYLVVPOQGFPCDPPAPGAGVMVHHRSSSTRSGGGLTLGLPEPSEBA PRSPLAPSEG 1080  
 Qy 1081 AGSDVPEGDLGMGAAGKLOSLPTHDSPLQRYSEDPVLPSETDGYVAPLTCSPQBEYV 1140  
 Db 1081 AGSDVPEGDLGMGAAGKLOSLPTHDSPLQRYSEDPVLPSETDGYVAPLTCSPQBEYV 1140  
 Qy 1141 NOPVRRPQPSRRGGLPAARPAAGATLERPKTLSPGKGVVVDVAFAGGAIVENBEYLTPO 1200  
 Db 1141 NOPVRRPQPSRRGGLPAARPAAGATLERPKTLSPGKGVVVDVAFAGGAIVENBEYLTPO 1200  
 Qy 1201 GGAAPQHPPPAPFPAFNDLYWMDQDPPERGAPSTFKGPTAENPEYLGIDVYV 1255  
 Db 1201 GGAAPQHPPPAPFPAFNDLYWMDQDPPERGAPSTFKGPTAENPEYLGIDVYV 1255

RESULT 7  
 US-08-486-348A-68  
 ; Sequence 68, Application US/08486348A  
 ; Patent No. 5846538

```

GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

Query Match 99.9%; Score 6806; DB 1; Length 1255;
Beet local similarity 99.9%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

481 PWDQFRNPHQALHTANRPEDECEVGEGLACHOLCARGHCWGPPTQCVCNSQFLRGEC 540
481 PWDQFRNPHQALHTANRPEDECEVGEGLACHOLCARGHCWGPPTQCVCNSQFLRGEC 540
541 VEECKVTLQGLPREYVNAHCLPCHECOPQNGSVTCFEPADQCACAHYDPPFCVARC 600
541 VEECKVTLQGLPREYVNAHCLPCHECOPQNGSVTCFEPADQCACAHYDPPFCVARC 600
601 PSQVPELSTYMPIMKPEPBEBCOPCPINCCHSCVLDLKDCKCPABORASPLTISAYVG 660
601 PSQVPELSTYMPIMKPEPBEBCOPCPINCCHSCVLDLKDCKCPABORASPLTISAYVG 660
661 ILVVVLGVVFGILIKRQOKIRKXTMRLLQETELVEPLTPSGAMPQAOQRIKTEL 720
661 ILVVVLGVVFGILIKRQOKIRKXTMRLLQETELVEPLTPSGAMPQAOQRIKTEL 720
721 RKVYLSGAFGTYYKGIWIPDGENVKIPVAIKVIRENTSRKANKELIDEAYVAGVSP 780
721 RKVYLSGAFGTYYKGIWIPDGENVKIPVAIKVIRENTSRKANKELIDEAYVAGVSP 780
781 YVSRLLGI CLSTVQVLTQMLMPYGLADHVENRGRGLSODLLMCMQIAGMSYLDVR 840
781 YVSRLLGI CLSTVQVLTQMLMPYGLADHVENRGRGLSODLLMCMQIAGMSYLDVR 840
841 LVHRDLAARNVLYKSPNHVKITDFGLARLDIDETRYHADGAKVPIKMMALESILRRFT 900
841 LVHRDLAARNVLYKSPNHVKITDFGLARLDIDETRYHADGAKVPIKMMALESILRRFT 900
901 HOSDWSYGVTVWEIMTFGAKPYDGI PARBIPDLLEKBERLPPICTIDVYMIWYKCM 960
901 HOSDWSYGVTVWEIMTFGAKPYDGI PARBIPDLLEKBERLPPICTIDVYMIWYKCM 960
961 IDSECRPRFRELVSFESPMARDPQRFVYIQNBDLSPASPLDSTFRSLDEDDMDGLVDA 1020
961 IDSECRPRFRELVSFESPMARDPQRFVYIQNBDLSPASPLDSTFRSLDEDDMDGLVDA 1020
1021 BEYLVPOQGFPCPPAPAGGVVHHRHS$STRSGGDLTLGLBPSBEBAPRSLAPSEG 1080
1021 BEYLVPOQGFPCPPAPAGGVVHHRHS$STRSGGDLTLGLBPSBEBAPRSLAPSEG 1080
1081 AGSDVFDGDLGMAKGIQSLPTHDSPQLQRYSEDPYPLPSETDGYVAPLTCSPQPEYV 1140
1081 AGSDVFDGDLGMAKGIQSLPTHDSPQLQRYSEDPYPLPSETDGYVAPLTCSPQPEYV 1140
1141 NOPDVRPQPPSPREBPLPAABPAGATTLERPTLSGKGVYKDVAFGAVENPEYLTPO 1200
1141 NOPDVRPQPPSPREBPLPAABPAGATTLERPTLSGKGVYKDVAFGAVENPEYLTPO 1200
1201 GGAAPQPPPPAFSPAFDNLVYMDODPPERKAPSTPFGKTPAENPEYLGIDVPV 1255
1201 GGAAPQPPPPAFSPAFDNLVYMDODPPERKAPSTPFGKTPAENPEYLGIDVPV 1255

RESULT 8
US-08-486-545B-68
Sequence 68, Application US/08468545B
Patent No. 5876712
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,545B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-468-545B-68

```

```

Query Match      99.9%; Score 6806; DB 1; Length 1255;
Beet Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;
Matches 1254; Conservative 0;

```

```

OY 1 MELAALCRMGILLALLPPGAASVQVCTGTDMKRLPASPEITHLDMRLHLYGSCQVYVQNL 60
DB 1 MELAALCRMGILLALLPPGAASVQVCTGTDMKRLPASPEITHLDMRLHLYGSCQVYVQNL 60
OY 61 ELTYIPTNASISFLQDIOBVQGYVLIANHVRQVPLQRLRYRGTLQFEDNYALAVLDNG 120
DB 61 ELTYIPTNASISFLQDIOBVQGYVLIANHVRQVPLQRLRYRGTLQFEDNYALAVLDNG 120
OY 121 DPLNNTTAVTASPGSLRELQRLSLTEILKGGVLIQRNPOLCYQDTIIMKQIFKNNOLA 180
DB 121 DPLNNTTAVTASPGSLRELQRLSLTEILKGGVLIQRNPOLCYQDTIIMKQIFKNNOLA 180
OY 181 LTLIDTNSRACHPCSPMCKSGRCWGESSEDCQSLTRTVACGACGRCARCKPLPTDCHEQC 240
DB 181 LTLIDTNSRACHPCSPMCKSGRCWGESSEDCQSLTRTVACGACGRCARCKPLPTDCHEQC 240
OY 241 AAGCTGPRHSDCLALHFNHSGICELHCPALVTYNTDFESMPNREGRTYTGASCVTACP 300
DB 241 AAGCTGPRHSDCLALHFNHSGICELHCPALVTYNTDFESMPNREGRTYTGASCVTACP 300
OY 301 YNYLSTDVSGCTLVCPLNQEVTAEDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
DB 301 YNYLSTDVSGCTLVCPLNQEVTAEDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
OY 361 IOEFAGCKKIFGSLAFLPESPDGASNTAPLQPEQLQVFTLEBITGYLYISAMPDSL 420
DB 361 IOEFAGCKKIFGSLAFLPESPDGASNTAPLQPEQLQVFTLEBITGYLYISAMPDSL 420
OY 421 DLSVQNTQVIRGRILIANGAYSLTQIGISWLGIRSLRELSGIALIHNTHLCFVHTV 480
DB 421 DLSVQNTQVIRGRILIANGAYSLTQIGISWLGIRSLRELSGIALIHNTHLCFVHTV 480
OY 481 PMDQFRPHQALHTANRPEBCEVSEGGLAQHQLCARGHCKMPGTQCVNCSQPLRGSEC 540
DB 481 PMDQFRPHQALHTANRPEBCEVSEGGLAQHQLCARGHCKMPGTQCVNCSQPLRGSEC 540
OY 541 VEECVLQGLPREVYNARHCLPCHPECOPONGSVTCFGEADQCVACAHYKDPPECVAR 600
DB 541 VEECVLQGLPREVYNARHCLPCHPECOPONGSVTCFGEADQCVACAHYKDPPECVAR 600
OY 601 PSGVPRDLSYMPITWKFPEBEGAQCPQPINCTHSQVLDLDDKGPABQASPLTSISAVG 660
DB 601 PSGVPRDLSYMPITWKFPEBEGAQCPQPINCTHSQVLDLDDKGPABQASPLTSISAVG 660
OY 661 ILLVVVLGVIRGILIKRQOKIRKTYMRLLOETELVPLPSSGAMPQAOIRLIKETEL 720
DB 661 ILLVVVLGVIRGILIKRQOKIRKTYMRLLOETELVPLPSSGAMPQAOIRLIKETEL 720

```

```

OY 721 RKVAVLGSAGFVYKGIWIPDGENVKI PVAIKYLRENTSPEKANEIIDEAYYVAGVSP 780
DB 721 RKVAVLGSAGFVYKGIWIPDGENVKI PVAIKYLRENTSPEKANEIIDEAYYVAGVSP 780
OY 781 YVSRLLGICLSTVQVLTQMLPYGCLLDHVENNGRLGSDOLMCMQOIAAGMSLBEVR 840
DB 781 YVSRLLGICLSTVQVLTQMLPYGCLLDHVENNGRLGSDOLMCMQOIAAGMSLBEVR 840
OY 841 LVHRDLAARNLVKSPNHVKITDGLARLADIDETEVHADGKVPKIMMALESILRRFT 900
DB 841 LVHRDLAARNLVKSPNHVKITDGLARLADIDETEVHADGKVPKIMMALESILRRFT 900
OY 901 HQSDVMSYGVTVWELMTFPAKAYDGIIPAREIPDLLEKGERLPOPPICITDYYMTMKCM 960
DB 901 HQSDVMSYGVTVWELMTFPAKAYDGIIPAREIPDLLEKGERLPOPPICITDYYMTMKCM 960
OY 961 IDSECRPRFRELVEFSRMAKDPPRVITQNEDEDGPSPLDSTYRSLLBEDDMDGLVDA 1020
DB 961 IDSECRPRFRELVEFSRMAKDPPRVITQNEDEDGPSPLDSTYRSLLBEDDMDGLVDA 1020
OY 1021 EEYLVPOQGFPCPDPAAGAGMVRHRSSTRSGGDLTLGLPSEBEARSPPLAPSEG 1080
DB 1021 EEYLVPOQGFPCPDPAAGAGMVRHRSSTRSGGDLTLGLPSEBEARSPPLAPSEG 1080
OY 1081 AGSDVFDGLMGAKGIQSLPTHDPSPLOKYSDEPTVPLPSEYDGYVALTCSPOPEYV 1140
DB 1081 AGSDVFDGLMGAKGIQSLPTHDPSPLOKYSDEPTVPLPSEYDGYVALTCSPOPEYV 1140
OY 1141 NOPVRPOPSPRBRGPIPAARPAATLBRPTLSPGKNGVVDPAFGAAYENPEYLTPO 1200
DB 1141 NOPVRPOPSPRBRGPIPAARPAATLBRPTLSPGKNGVVDPAFGAAYENPEYLTPO 1200
OY 1201 GAAPQHPHPPAFDNLVYMDQDPPERGAPSTFGPTAENPEYLGIDVY 1255
DB 1201 GAAPQHPHPPAFDNLVYMDQDPPERGAPSTFGPTAENPEYLGIDVY 1255

```

```

RESULT 9
US-08-466-680B-68
; Sequence 68, Application US/08466680B
; Patent No. 6075122
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disls, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,680B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 1255 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-08-466-680B-68

Query Match 99.9%; Score 6806; DB 2; Length 1255;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

1  MELAALCRKGLLALLPFGAASQVCTGDMKRLRLPASBETHLDMRLHYOCQVQGNL 60
1  MELAALCRKGLLALLPFGAASQVCTGDMKRLRLPASBETHLDMRLHYOCQVQGNL 60
1  MELAALCRKGLLALLPFGAASQVCTGDMKRLRLPASBETHLDMRLHYOCQVQGNL 60
61  ELYLPTNASLSFLQDIOEVQGYVLIANHQRVPLQRLIRVGTQLFEDNYALAVLNG 120
61  ELYLPTNASLSFLQDIOEVQGYVLIANHQRVPLQRLIRVGTQLFEDNYALAVLNG 120
61  ELYLPTNASLSFLQDIOEVQGYVLIANHQRVPLQRLIRVGTQLFEDNYALAVLNG 120
121  DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTILMKDIFHKNNQLA 180
121  DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTILMKDIFHKNNQLA 180
121  DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTILMKDIFHKNNQLA 180
181  LTLIDTNRBRACHPCSPMKSGSRGWESSEDCOSLTRTYCAGGCARCKGPLPTDCHEQC 240
181  LTLIDTNRBRACHPCSPMKSGSRGWESSEDCOSLTRTYCAGGCARCKGPLPTDCHEQC 240
181  LTLIDTNRBRACHPCSPMKSGSRGWESSEDCOSLTRTYCAGGCARCKGPLPTDCHEQC 240
241  AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTYFGASCVTACP 300
241  AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTYFGASCVTACP 300
241  AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTYFGASCVTACP 300
301  YNYLSTDVSGCTLVCPPLHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
301  YNYLSTDVSGCTLVCPPLHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
301  YNYLSTDVSGCTLVCPPLHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
361  IOEFGAGCKKIFGSLAFLESFPGDPASNTAPLOPBOLOVFETLEETITGLYISAMPDILP 420
361  IOEFGAGCKKIFGSLAFLESFPGDPASNTAPLOPBOLOVFETLEETITGLYISAMPDILP 420
361  IOEFGAGCKKIFGSLAFLESFPGDPASNTAPLOPBOLOVFETLEETITGLYISAMPDILP 420
421  DLSVFQNTQVIRGRILHNGAYSLLTQGLISWLGRLSRLBELSGALIHNNTHLCFVHTV 480
421  DLSVFQNTQVIRGRILHNGAYSLLTQGLISWLGRLSRLBELSGALIHNNTHLCFVHTV 480
421  DLSVFQNTQVIRGRILHNGAYSLLTQGLISWLGRLSRLBELSGALIHNNTHLCFVHTV 480
481  PWDOLFERNHQAALHNTANRPEDECVBGLACHQLCARGHCMWGPPTQVNCSSQPLRGQBC 540
481  PWDOLFERNHQAALHNTANRPEDECVBGLACHQLCARGHCMWGPPTQVNCSSQPLRGQBC 540
481  PWDOLFERNHQAALHNTANRPEDECVBGLACHQLCARGHCMWGPPTQVNCSSQPLRGQBC 540
541  VEECEVLOGLPREYVNAARHCLPCHPECOQNGSVTCFGEADQVACAHYKDPFVCVAC 600
541  VEECEVLOGLPREYVNAARHCLPCHPECOQNGSVTCFGEADQVACAHYKDPFVCVAC 600
541  VEECEVLOGLPREYVNAARHCLPCHPECOQNGSVTCFGEADQVACAHYKDPFVCVAC 600
601  PEGVNPDLISYMPIMKRPDEEGAOCPCPICTHSQVLDLDDKGPABEGRASPLSIISAVVG 660
601  PEGVNPDLISYMPIMKRPDEEGAOCPCPICTHSQVLDLDDKGPABEGRASPLSIISAVVG 660
601  PEGVNPDLISYMPIMKRPDEEGAOCPCPICTHSQVLDLDDKGPABEGRASPLSIISAVVG 660
661  ILLVVLGVVFGILIKRROOKIRKTYMRLLQETELVEPLTPSGAMPQNAQMRILKETEL 720
661  ILLVVLGVVFGILIKRROOKIRKTYMRLLQETELVEPLTPSGAMPQNAQMRILKETEL 720
661  ILLVVLGVVFGILIKRROOKIRKTYMRLLQETELVEPLTPSGAMPQNAQMRILKETEL 720
721  RKRYVLGSGAGFTVYKGIWIPDGENYKIPVAIKVIRENTSPRANKIILDEAYVMAGVGP 780
721  RKRYVLGSGAGFTVYKGIWIPDGENYKIPVAIKVIRENTSPRANKIILDEAYVMAGVGP 780
721  RKRYVLGSGAGFTVYKGIWIPDGENYKIPVAIKVIRENTSPRANKIILDEAYVMAGVGP 780
781  VYSRLIGICLSTVQVLTQVLMFYGCLLDHVRNRLGSLGODLLNMCQIAKMSYLEDVR 840
781  VYSRLIGICLSTVQVLTQVLMFYGCLLDHVRNRLGSLGODLLNMCQIAKMSYLEDVR 840
781  VYSRLIGICLSTVQVLTQVLMFYGCLLDHVRNRLGSLGODLLNMCQIAKMSYLEDVR 840
841  LVHRDLAARNVAVKSPNHVYKTDVGLARLLDIDETRYHADGKVPILKMMALSLIRRRPT 900
841  LVHRDLAARNVAVKSPNHVYKTDVGLARLLDIDETRYHADGKVPILKMMALSLIRRRPT 900
841  LVHRDLAARNVAVKSPNHVYKTDVGLARLLDIDETRYHADGKVPILKMMALSLIRRRPT 900
901  HOSDWSYGVYTWELMTFGAKPYDGI PAEIPDLLEKGRLLPOPPICITIDVYMIWYKCM 960
901  HOSDWSYGVYTWELMTFGAKPYDGI PAEIPDLLEKGRLLPOPPICITIDVYMIWYKCM 960
901  HOSDWSYGVYTWELMTFGAKPYDGI PAEIPDLLEKGRLLPOPPICITIDVYMIWYKCM 960
961  DLSVFQNTQVIRGRILHNGAYSLLTQGLISWLGRLSRLBELSGALIHNNTHLCFVHTV 480
961  DLSVFQNTQVIRGRILHNGAYSLLTQGLISWLGRLSRLBELSGALIHNNTHLCFVHTV 480
961  DLSVFQNTQVIRGRILHNGAYSLLTQGLISWLGRLSRLBELSGALIHNNTHLCFVHTV 480

```

Query Match 99.9%; Score 6806; DB 2; Length 1255;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

961  IDSECRPRERELVSFBSRMARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDDMGDLVDA 1020
961  IDSECRPRERELVSFBSRMARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDDMGDLVDA 1020
961  IDSECRPRERELVSFBSRMARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDDMGDLVDA 1020
1021  EBYLVPQCGFCPPDPAPGAGVHHRSSSTRSGGDLTGLSPSEEAAPSPPLAPSG 1080
1021  EBYLVPQCGFCPPDPAPGAGVHHRSSSTRSGGDLTGLSPSEEAAPSPPLAPSG 1080
1021  EBYLVPQCGFCPPDPAPGAGVHHRSSSTRSGGDLTGLSPSEEAAPSPPLAPSG 1080
1081  AGSDVPDGLGMAKGLQSLPTHPSPLOQRYSBPVPLPSETQGYVAPLCSPOPEYV 1140
1081  AGSDVPDGLGMAKGLQSLPTHPSPLOQRYSBPVPLPSETQGYVAPLCSPOPEYV 1140
1081  AGSDVPDGLGMAKGLQSLPTHPSPLOQRYSBPVPLPSETQGYVAPLCSPOPEYV 1140
1141  NOPDVRPOPSPREBGLPAPAPAGATLERPKTLSPGKGVVDVPAFGAVENPEYLTQ 1200
1141  NOPDVRPOPSPREBGLPAPAPAGATLERPKTLSPGKGVVDVPAFGAVENPEYLTQ 1200
1141  NOPDVRPOPSPREBGLPAPAPAGATLERPKTLSPGKGVVDVPAFGAVENPEYLTQ 1200
1201  GGAAPQHPHPAPSPAFDNLXYWDODPPERKAPPESTFKCTPFAENPEYLGDLVPV 1255
1201  GGAAPQHPHPAPSPAFDNLXYWDODPPERKAPPESTFKCTPFAENPEYLGDLVPV 1255
1201  GGAAPQHPHPAPSPAFDNLXYWDODPPERKAPPESTFKCTPFAENPEYLGDLVPV 1255

```

RESULT 10  
 US-09-527-487-2  
 ; Sequence 2, Application US/09527487  
 ; Patent No. 6528060  
 ; APPLICANT: Nicolette, Charles  
 ; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES  
 ; FILE REFERENCE: 126881309200  
 ; CURRENT APPLICATION NUMBER: US/09/527,487  
 ; CURRENT FILING DATE: 2000-03-16  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 1255  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-527-487-2

```
Db 421 DLSVFQNIQVIRGRILHNGAVSLTLQGLGISWLGRLSRLNLSGLALIHNTHTLCFVHTV 480
Qy 461 PMDQLFRNPHQALHTANRPEDECVGEGLACHQLCARGHCMGPGPTQCVNCSQFLRGEC 540
Db 461 PMDQLFRNPHQALHTANRPEDECVGEGLACHQLCARGHCMGPGPTQCVNCSQFLRGEC 540
Qy 541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPFCVARC 600
Db 541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPFCVARC 600
Qy 601 PSGVPELSTYMPIMKFPDEEGACOPCPINCTHSCVDLDDKCCPAEQRASPLTSISAVVG 660
Db 601 PSGVPELSTYMPIMKFPDEEGACOPCPINCTHSCVDLDDKCCPAEQRASPLTSISAVVG 660
Qy 661 ILVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOQMRILKETEL 720
Db 661 ILVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOQMRILKETEL 720
Qy 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELIDEAYVMAGVSP 780
Db 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELIDEAYVMAGVSP 780
Qy 781 YVSRLLGLCLTSTVOLVQLMPYGCILDHVENRGRLSQDLNMCQIAGKMSYLEBVR 840
Db 781 YVSRLLGLCLTSTVOLVQLMPYGCILDHVENRGRLSQDLNMCQIAGKMSYLEBVR 840
Qy 841 LVHRDLAARNVLVKSPPNHVKITDFGLARLLIDETEVHADGKVPDKMMALESILRRFT 900
Db 841 LVHRDLAARNVLVKSPPNHVKITDFGLARLLIDETEVHADGKVPDKMMALESILRRFT 900
Qy 901 HQSDWMSGVTVWELMTFGAKPYDGI PARFIPDLLEKGRLLPQPICTIDVYMIWVKCM 960
Db 901 HQSDWMSGVTVWELMTFGAKPYDGI PARFIPDLLEKGRLLPQPICTIDVYMIWVKCM 960
Qy 961 IDSECRPRFRELVSFESRMARDPQRFVVIQNEDELGPASPLDSTFYRSLLEDDMGDLVDA 1020
Db 961 IDSECRPRFRELVSFESRMARDPQRFVVIQNEDELGPASPLDSTFYRSLLEDDMGDLVDA 1020
Qy 1021 EBYLVPQOGFPCPDPAFGAGVHHRHSSSTRSGGDLTLGLBESSEBAPRSLAPSEG 1080
Db 1021 EBYLVPQOGFPCPDPAFGAGVHHRHSSSTRSGGDLTLGLBESSEBAPRSLAPSEG 1080
Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
Qy 1141 NOPDVRPQPSRREGPLPAARPAAGATLERAKTSLSGKGVKDVAPAGAVENPEYLTPO 1200
Db 1141 NOPDVRPQPSRREGPLPAARPAAGATLERAKTSLSGKGVKDVAPAGAVENPEYLTPO 1200
Qy 1201 GGAAPQPHPPAFSPAFDNLVYWDQPPERGAPESTFGKTPTAENPEYLGLDVPI 1255
Db 1201 GGAAPQPHPPAFSPAFDNLVYWDQPPERGAPESTFGKTPTAENPEYLGLDVPI 1255

RESULT 11
US-09-811-115-3
; Sequence 3, Application us/09811115
; Parent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENE 034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1255
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3
Query Match 99.8%; Score 6806; DB 2; Length 1255;
Best Local Similarity 99.8%; Fred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELAALCMWGLLMLALPPGAASSTQVCTGDMKRLRPAPETHLIDLRLHYOGCCQVVGNTL 60
Db 1 MELAALCMWGLLMLALPPGAASSTQVCTGDMKRLRPAPETHLIDLRLHYOGCCQVVGNTL 60
Qy 61 ELTYLPTNASISPLQDIOEVQGVYLIANQVROVPLQRLRYRGTQLFEDNYALAVLNG 120
Db 61 ELTYLPTNASISPLQDIOEVQGVYLIANQVROVPLQRLRYRGTQLFEDNYALAVLNG 120
Qy 121 DPLNTPVWASPGGLRELOLRSTETLKGVLIQRPOLCYQDTILMKDIFKNNOLA 180
Db 121 DPLNTPVWASPGGLRELOLRSTETLKGVLIQRPOLCYQDTILMKDIFKNNOLA 180
Qy 181 LTLIDTNRSRACHPCSPMKSGSRGESSBDCQSLTRTVCAAGCARCKGPLPTDCCHQC 240
Db 181 LTLIDTNRSRACHPCSPMKSGSRGESSBDCQSLTRTVCAAGCARCKGPLPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNDEGYRTGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNDEGYRTGASCVTACP 300
Qy 301 YNYLSTDVSGCTVACPLHNOEVTAE DGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
Db 301 YNYLSTDVSGCTVACPLHNOEVTAE DGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
Qy 361 IOEFAGCKKIFGSLAFUPRESFDGPASNTABLOEQLOVFETLBEITGYLIISAMPDLSL 420
Db 361 IOEFAGCKKIFGSLAFUPRESFDGPASNTABLOEQLOVFETLBEITGYLIISAMPDLSL 420
Qy 421 DLSVFQNIQVIRGRILHNGAVSLTLQGLGISWLGRLSRLNLSGLALIHNTHTLCFVHTV 480
Db 421 DLSVFQNIQVIRGRILHNGAVSLTLQGLGISWLGRLSRLNLSGLALIHNTHTLCFVHTV 480
Qy 481 PMDQLFRNPHQALHTANRPEDECVGEGLACHQLCARGHCMGPGPTQCVNCSQFLRGEC 540
Db 481 PMDQLFRNPHQALHTANRPEDECVGEGLACHQLCARGHCMGPGPTQCVNCSQFLRGEC 540
Qy 541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPFCVARC 600
Db 541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPFCVARC 600
Qy 601 PSGVPELSTYMPIMKFPDEEGACOPCPINCTHSCVDLDDKCCPAEQRASPLTSISAVVG 660
Db 601 PSGVPELSTYMPIMKFPDEEGACOPCPINCTHSCVDLDDKCCPAEQRASPLTSISAVVG 660
Qy 661 ILVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOQMRILKETEL 720
Db 661 ILVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOQMRILKETEL 720
Qy 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELIDEAYVMAGVSP 780
Db 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPKANKELIDEAYVMAGVSP 780
Qy 781 YVSRLLGLCLTSTVOLVQLMPYGCILDHVENRGRLSQDLNMCQIAGKMSYLEBVR 840
Db 781 YVSRLLGLCLTSTVOLVQLMPYGCILDHVENRGRLSQDLNMCQIAGKMSYLEBVR 840
Qy 841 LVHRDLAARNVLVKSPPNHVKITDFGLARLLIDETEVHADGKVPDKMMALESILRRFT 900
Db 841 LVHRDLAARNVLVKSPPNHVKITDFGLARLLIDETEVHADGKVPDKMMALESILRRFT 900
Qy 901 HQSDWMSGVTVWELMTFGAKPYDGI PARFIPDLLEKGRLLPQPICTIDVYMIWVKCM 960
Db 901 HQSDWMSGVTVWELMTFGAKPYDGI PARFIPDLLEKGRLLPQPICTIDVYMIWVKCM 960
Qy 961 IDSECRPRFRELVSFESRMARDPQRFVVIQNEDELGPASPLDSTFYRSLLEDDMGDLVDA 1020
```

```

Db      961  IDSECRPRFRELVSFSRWARDPQRFVYIQNEDLCPASPLDSTFYSRLLEDMDMDLVDVA 1020
Qy      1021  EBYLVPOGFCPPDPAPGAGVWHRHRSSTRSGGDTTLGLBSEBEPAPSLAPSBG 1080
Db      1021  EBYLVPOGFCPPDPAPGAGVWHRHRSSTRSGGDTTLGLBSEBEPAPSLAPSBG 1080
Qy      1081  AGSDVFDGDLGMAKAGLQSLPTHDPSLQRYSEDPVPLPSETDGYAAPLTCSPQPEVY 1140
Db      1081  AGSDVFDGDLGMAKAGLQSLPTHDPSLQRYSEDPVPLPSETDGYAAPLTCSPQPEVY 1140
Qy      1141  NOPDVRPOPSPREBEPPLPAAPAGATLERPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
Db      1141  NOPDVRPOPSPREBEPPLPAAPAGATLERPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
Qy      1201  GGAAPQHPPPAFSPAFDNLTYWDODPPERGAPESTFKGTPTAENPEYLGIDVPV 1255
Db      1201  GGAAPQHPPPAFSPAFDNLTYWDODPPERGAPESTFKGTPTAENPEYLGIDVPV 1255

RESULT 12
US-09-354-533-68
; Sequence 68, Application US/09354533
; Patent No. 6664370
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; Diele, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/354,533
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-09-354-533-68

Query Match 99.9%; Score 6806; DB 2; Length 1255;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1554; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      121  DPLANTTQVTVGASPGGLRELQRLSLTBILKGVLIQRNPOLCYOPTILMKOIFHNQOLA 180
Db      121  DPLANTTQVTVGASPGGLRELQRLSLTBILKGVLIQRNPOLCYOPTILMKOIFHNQOLA 180
Qy      181  LTLIDTNRSRACHPCSPCKSGRCMGSESDCQSLTRTVCAAGCARCKGPLETDCCHQC 240
Db      181  LTLIDTNRSRACHPCSPCKSGRCMGSESDCQSLTRTVCAAGCARCKGPLETDCCHQC 240
Qy      241  AAGCTGPRGSDCLACLHFNHSGICELHCPALVTYNTDFESMPNREGRTFCAACVTACP 300
Db      241  AAGCTGPRGSDCLACLHFNHSGICELHCPALVTYNTDFESMPNREGRTFCAACVTACP 300
Qy      301  YNLTSDVSGCTLVCPLEHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREAVTSAN 360
Db      301  YNLTSDVSGCTLVCPLEHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREAVTSAN 360
Qy      361  IOEFAGCKKIFGSLAFPLESPFDGDBASNTAPLQPEBOLQVFTLEETGYLISAMPDSL 420
Db      361  IOEFAGCKKIFGSLAFPLESPFDGDBASNTAPLQPEBOLQVFTLEETGYLISAMPDSL 420
Qy      421  DLSVFQNTQVTRGRILNHGAYSLLTQIGISWLGSRSLRELGSGLALIHNTHLCEVHTV 480
Db      421  DLSVFQNTQVTRGRILNHGAYSLLTQIGISWLGSRSLRELGSGLALIHNTHLCEVHTV 480
Qy      481  PMDOLFRPHOALHTANRPEDECVGEGIAHQLCARGHCMGPGFTQCVCNCSQFLRGQC 540
Db      481  PMDOLFRPHOALHTANRPEDECVGEGIAHQLCARGHCMGPGFTQCVCNCSQFLRGQC 540
Qy      541  VEBECVLOGLPREYVNAHCLPCHBECOPQNGSVTCFGEADQCYACAHYKPPFCVARC 600
Db      541  VEBECVLOGLPREYVNAHCLPCHBECOPQNGSVTCFGEADQCYACAHYKPPFCVARC 600
Qy      601  PSGVPRDLSYMPYKFPPEBEGACOPCPINCHTSCVDLDDKGRAPORASPLTSIIISAVG 660
Db      601  PSGVPRDLSYMPYKFPPEBEGACOPCPINCHTSCVDLDDKGRAPORASPLTSIIISAVG 660
Qy      661  ILVVVLGVFGLILKRRQOKIRKTYMRLLQETELVEPLFPGSGAMPQAOQRILKETEL 720
Db      661  ILVVVLGVFGLILKRRQOKIRKTYMRLLQETELVEPLFPGSGAMPQAOQRILKETEL 720
Qy      721  RKVKVLGSGAFQTVYKGIWIPDGENVKI PVAKVLRNTSPKANKEIIDEAYVMAVGSP 780
Db      721  RKVKVLGSGAFQTVYKGIWIPDGENVKI PVAKVLRNTSPKANKEIIDEAYVMAVGSP 780
Qy      781  YVSRLLGICLTSTVOLVTLQMPYGCILDHVENRGLSODILNMCQIAKMSYLEDVR 840
Db      781  YVSRLLGICLTSTVOLVTLQMPYGCILDHVENRGLSODILNMCQIAKMSYLEDVR 840
Qy      841  LVHRDLAARNVLVYSPNHYKITDPGLARLDIDETEHADGKVPDKMMALESIIRRP 900
Db      841  LVHRDLAARNVLVYSPNHYKITDPGLARLDIDETEHADGKVPDKMMALESIIRRP 900
Qy      901  HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKBERLPQPICTIDVYMIWCKM 960
Db      901  HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKBERLPQPICTIDVYMIWCKM 960
Qy      961  IDSECRPRFRELVSFSRWARDPQRFVYIQNEDLCPASPLDSTFYSRLLEDMDMDLVDVA 1020
Db      961  IDSECRPRFRELVSFSRWARDPQRFVYIQNEDLCPASPLDSTFYSRLLEDMDMDLVDVA 1020
Qy      1021  EBYLVPOGFCPPDPAPGAGVWHRHRSSTRSGGDTTLGLBSEBEPAPSLAPSBG 1080
Db      1021  EBYLVPOGFCPPDPAPGAGVWHRHRSSTRSGGDTTLGLBSEBEPAPSLAPSBG 1080
Qy      1081  AGSDVFDGDLGMAKAGLQSLPTHDPSLQRYSEDPVPLPSETDGYAAPLTCSPQPEVY 1140
Db      1081  AGSDVFDGDLGMAKAGLQSLPTHDPSLQRYSEDPVPLPSETDGYAAPLTCSPQPEVY 1140
Qy      1141  NOPDVRPOPSPREBEPPLPAAPAGATLERPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
Db      1141  NOPDVRPOPSPREBEPPLPAAPAGATLERPKTLSPGKGVVDVAFGAVENPEYLTPO 1200

```

Qy 1201 GGAAPQHPHPAPSPAFNLYWMDPPERGAPESTFKGTPLEANPEYLGLDVPY 1255  
Db 1201 GGAAPQHPHPAPSPAFNLYWMDPPERGAPESTFKGTPLEANPEYLGLDVPY 1255

## RESULT 13

US-08-484-438-8

Sequence 8, Application US/08484438

Patent No. 5811098

Patent No. 5811098 5780031

## GENERAL INFORMATION:

APPLICANT: Plozman, Gregory D.

APPLICANT: Culouscou, Jean-Michel

APPLICANT: Shoyab, Mohammed

APPLICANT: Siegal, Clay B.

APPLICANT: Helicstr m, Ingegerd

APPLICANT: Helicstr m, Karl E.

TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie &amp; Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,438

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/323,442

FILING DATE: 14-OCT-1994

APPLICATION NUMBER: US 08/150,704

FILING DATE: 10-NOV-1993

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/981,165

FILING DATE: 24-NOV-1992

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Mastrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 5624-230

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-484-438-8

Query Match 98.8%; Score 6732; DB 1; Length 1255;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1447; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

Qy 1 MELALICRMGLLALLPFGAASSTOVCTGTMKRLRPASPTHLDMLRHLYOGCCQVVGNTL 60  
Db 1 MELALICRMGLLALLPFGAASSTOVCTGTMKRLRPASPTHLDMLRHLYOGCCQVVGNTL 60  
Qy 61 ELTYLPTNASLFLDDIOGVGCVLIANQVAVQLQRLRYRGQLTFEDNYALAVLDNG 120  
Db 61 ELTYLPTNASLFLDDIOGVGCVLIANQVAVQLQRLRYRGQLTFEDNYALAVLDNG 120  
Qy 61 ELTYLPTNASLFLDDIOGVGCVLIANQVAVQLQRLRYRGQLTFEDNYALAVLDNG 120  
Db 61 ELTYLPTNASLFLDDIOGVGCVLIANQVAVQLQRLRYRGQLTFEDNYALAVLDNG 120

Qy 121 DPLNNTPTVGTASPGGLRELQLRSLEILKGGVLIQBNPOLCYODTILMKDI FHKNNQLA 180  
Db 121 DPLNNTPTVGTASPGGLRELQLRSLEILKGGVLIQBNPOLCYODTILMKDI FHKNNQLA 180  
Qy 181 LTLIDTNRSRACHCSCKSGSRCKWSSSEDCOSLTTRVCAAGGACRCGRLPTCCCHQC 240  
Db 181 LTLIDTNRSRACHCSCKSGSRCKWSSSEDCOSLTTRVCAAGGACRCGRLPTCCCHQC 240  
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 300  
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 300  
Qy 301 YNYLSTDVGSCTTLCPLHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360  
Db 301 YNYLSTDVGSCTTLCPLHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360  
Qy 361 IOEFAGCKKIFGSLAFIPESFDGPAANTAPLOEQLOVFTLEITGYLISAMPISLP 420  
Db 361 IOEFAGCKKIFGSLAFIPESFDGPAANTAPLOEQLOVFTLEITGYLISAMPISLP 420  
Qy 421 DLSVFQNLQVYRGRILHNGAYSLTLOGLGISWLGRLSRLRELSGLALIHNTTHCFVYTV 480  
Db 421 DLSVFQNLQVYRGRILHNGAYSLTLOGLGISWLGRLSRLRELSGLALIHNTTHCFVYTV 480  
Qy 481 PMDQLFNPNHQAALHTANRPEDECVGEGLAHQLCARHGCWGPPTQCVNCSQFLRGQC 540  
Db 481 PMDQLFNPNHQAALHTANRPEDECVGEGLAHQLCARHALLGSGPTQCVNCSQFLRGQC 540  
Qy 541 VEECRVYQGLPREVYVNAHCLPCHPECPQNGSTTCGPEADOCVACHYVDPFCVARC 600  
Db 541 VEECRVYQGLPREVYVNAHCLPCHPECPQNGSTTCGPEADOCVACHYVDPFCVARC 600  
Qy 601 PSQVKPDLSTYMPYKPFDEBACQPCPINCTHSCVDLDDKCPABQASPLTSLVSNVVG 660  
Db 601 PSQVKPDLSTYMPYKPFDEBACQPCPINCTHSCVDLDDKCPABQASPLTSLVSNVVG 660  
Qy 661 ILVAVVLGVVEGILIKRQOKIRKRYNRRLQETELVEPLTPSGAMPNOAMRLKETEL 720  
Db 661 ILVAVVLGVVEGILIKRQOKIRKRYNRRLQETELVEPLTPSGAMPNOAMRLKETEL 720  
Qy 721 RKVAVLGSSAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKAKETILDEAYVNAVGSP 780  
Db 721 RKVAVLGSSAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKAKETILDEAYVNAVGSP 780  
Qy 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYENRGRLSQDILLNMCQIAKMSYLEDVR 840  
Db 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYENRGRLSQDILLNMCQIAKMSYLEDVR 840  
Qy 841 LVHRDLAARNVLVKS PNHVKITDFGLARLLDIDETRYHADGKVPYIKMALESILRRFT 900  
Db 841 LVHRDLAARNVLVKS PNHVKITDFGLARLLDIDETRYHADGKVPYIKMALESILRRFT 900  
Qy 901 HQSDVWSGYTVWMLMTFGAKPYDGI PARETIDLLLEGGERLPORPICIDVYMMVKKM 960  
Db 901 HQSDVWSGYTVWMLMTFGAKPYDGI PARETIDLLLEGGERLPORPICIDVYMMVKKM 960  
Qy 961 IDSECRPPFRELVESEFARMADPQRFVYIQNEBGLPASPLDSTFYRSLLEDMDGLVDA 1020  
Db 961 IDSECRPPFRELVESEFARMADPQRFVYIQNEBGLPASPLDSTFYRSLLEDMDGLVDA 1020  
Qy 1021 EEYLVPQGFPCPDPAFGAGVHRRSSSTRSGGDLTLGLEPSEBEADRSPLAPSEG 1080  
Db 1021 EEYLVPQGFPCPDPAFGAGVHRRSSSTRSGGDLTLGLEPSEBEADRSPLAPSEG 1080  
Qy 1081 AGSVFPGDGLMGAAKGIQSLPTHDPSPLOQYSSDPVYPLPSEFDGYVAPLTCSPPQRYV 1140  
Db 1081 AGSVFPGDGLMGAAKGIQSLPTHDPSPLOQYSSDPVYPLPSEFDGYVAPLTCSPPQRYV 1140  
Qy 1141 NQPDVRPOPSPESRGSLPAAPAGATLERPTLSPGKGVVADYAFAGAVENPEYLTPO 1200  
Db 1141 NQPDVRPOPSPESRGSLPAAPAGATLERPTLSPGKGVVADYAFAGAVENPEYLTPO 1200



QY 1201 GGAAQPPHPPAFSPAFDNLVYMDODPPERGAPESTFKGTPF-AENPEYLGADVY 1255  
DB 1201 GGAAQPPHPPAFSPAFDNLVYMDODPPERGAPESTFKGTPF-AENPEYLGADVY 1255

RESULT 14  
US-09-146-283-4  
; Sequence 4, Application US/09146283  
; Patent No. 5976546  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Ruegg, Curtis L.  
; APPLICANT: Wu, Hongyu  
; TITLE OF INVENTION: Immunostimulatory Compositions  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSES: Dehlinger & Associates  
; STREET: 350 Cambridge Ave. Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/146,283  
; FILING DATE: 03-SEPT-1998  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Judge, Linda R.  
; REGISTRATION NUMBER: 42,702  
; REFERENCE/DOCKET NUMBER: 7636-0010.21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-0880  
; TELEFAX: 650-324-0960  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 782 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: homo sapiens  
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
; US-09-146-283-4

Query Match 53.3%; Score 3630; DB 1; Length 782;  
Best Local Similarity 99.8%; Pred. No. 5.9e-238;  
Matches 653; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELALCWMGULLALLPFGASTOVCTGDMKLRLPASPEYTHLMDLRHLYOCGVVQGL 60  
DB 1 MELALCWMGULLALLPFGASTOVCTGDMKLRLPASPEYTHLMDLRHLYOCGVVQGL 60

QY 61 ELLVPTNASLSFLDIOEVGVYLIANQVQVPLQRLIVRGTOLEFEDNYALVLDNG 120  
DB 61 ELLVPTNASLSFLDIOEVGVYLIANQVQVPLQRLIVRGTOLEFEDNYALVLDNG 120

QY 121 DPLNNTTPTVGASPGGLRELQRLSLTEIKGCVLIQRNPQLCYODTIIMKDIFFHKNOLA 180  
DB 121 DPLNNTTPTVGASPGGLRELQRLSLTEIKGCVLIQRNPQLCYODTIIMKDIFFHKNOLA 180

QY 181 LTLITDNRBRACHPSGPMCKGSRGWESSEDCOSLTRITVACAGCARCKGPLPTDCHEQC 240  
DB 181 LTLITDNRBRACHPSGPMCKGSRGWESSEDCOSLTRITVACAGCARCKGPLPTDCHEQC 240

QY 241 AAGCTGPKHSDCLACIHFHSGICELHCPALVTYNTDFESNPNBGRYTFGASCVTACP 300  
DB 241 AAGCTGPKHSDCLACIHFHSGICELHCPALVTYNTDFESNPNBGRYTFGASCVTACP 300

QY 301 YNYLSTDVSGCTLYVCPILHNOEVTAEEDGTORCEKSKPCARVYCGHMEHLREAVATVSAN 360  
DB 301 YNYLSTDVSGCTLYVCPILHNOEVTAEEDGTORCEKSKPCARVYCGHMEHLREAVATVSAN 360

QY 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPELOVFEETLEITGYLIISAMPDILP 420  
DB 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPELOVFEETLEITGYLIISAMPDILP 420

QY 421 DLSVFQNTQVIRGRILHNGASVLTQIGISWLGRLSRLBELASGLALIHNTHLCPVHTV 480  
DB 421 DLSVFQNTQVIRGRILHNGASVLTQIGISWLGRLSRLBELASGLALIHNTHLCPVHTV 480

QY 481 PMDOLFRNPHOALHTTANRPEDECVGEGIAHQOLCARGHCMGPGPTOCVCSQFLRGQEC 540  
DB 481 PMDOLFRNPHOALHTTANRPEDECVGEGIAHQOLCARGHCMGPGPTOCVCSQFLRGQEC 540

QY 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGBEADQVACAHYKDPFCVARC 600  
DB 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGBEADQVACAHYKDPFCVARC 600

QY 601 PSQVKEPDLSTYMPYKFPDEBGAQCPINCTHSCVDLDDKGCPRARASPLTSL 654  
DB 601 PSQVKEPDLSTYMPYKFPDEBGAQCPINCTHSCVDLDDKGCPRARASPLTSL 654

RESULT 15  
US-08-579-823A-4  
; Sequence 4, Application US/08579823A  
; Patent No. 6080409  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Ruegg, Curtis L.  
; APPLICANT: Wu, Hongyu  
; TITLE OF INVENTION: Immunostimulatory Composition and Method  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSES: Dehlinger & Associates  
; STREET: 350 Cambridge Ave. Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/579,823A  
; FILING DATE: 03-DEC-1998  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Judge, Linda R.  
; REGISTRATION NUMBER: 42,702  
; REFERENCE/DOCKET NUMBER: 7636-0010  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-0880  
; TELEFAX: 650-324-0960  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 782 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: homo sapiens  
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
; US-08-579-823A-4

Query Match 53.3%; Score 3630; DB 2; Length 782;  
Best Local Similarity 99.8%; Pred. No. 5.9e-238;

```

Matches 653; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELAALCRMGILLIALLPPGAASVCTGTDMKLRLPASPEETHLDMLRHLYOGCQVVGNTL 60
Db 1 MELAALCRMGILLIALLPPGAASVCTGTDMKLRLPASPEETHLDMLRHLYOGCQVVGNTL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGVYLIANHNVQVPLQRLRIVRGTOLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIOEVQGVYLIANHNVQVPLQRLRIVRGTOLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTIIMKDIFFHKNOLA 180
Db 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTIIMKDIFFHKNOLA 180
QY 181 LTLIDPTNSRACHPCSPMCKSGRCWGESSEDCOSLTRTVACAGCARCKGPLPTDCHEQC 240
Db 181 LTLIDPTNSRACHPCSPMCKSGRCWGESSEDCOSLTRTVACAGCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMNPDEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMNPDEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCEPLHNOEVTAEQTCRCKSKPCARVCYGLGMEHLREVRVTSAN 360
Db 301 YNYLSTDVGSCTLVCEPLHNOEVTAEQTCRCKSKPCARVCYGLGMEHLREVRVTSAN 360
QY 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEETGYLYISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEETGYLYISAMPDSL 420
QY 421 DLSVFQNTQVIRGRILHNGAVSLTQGLGISWLGRLSLRELSSGLAIHNTHTLCFVHTV 480
Db 421 DLSVFQNTQVIRGRILHNGAVSLTQGLGISWLGRLSLRELSSGLAIHNTHTLCFVHTV 480
QY 481 PMDQLFRRNPQALLHTANRPEDECVGEGIACHQLCARGHCWGPPTCCVNCSSQFLRGQEC 540
Db 481 PMDQLFRRNPQALLHTANRPEDECVGEGIACHQLCARGHCWGPPTCCVNCSSQFLRGQEC 540
QY 541 VEECEVLQGLPREYVNAHCHLPCHEPCQPONGSVTCFGEADQCVACAHYKDPFCVARC 600
Db 541 VEECEVLQGLPREYVNAHCHLPCHEPCQPONGSVTCFGEADQCVACAHYKDPFCVARC 600
QY 601 PSGVKPDLISYMPIWKFPDEEGACQPCPINCTHSCVDLDDKGCAPAEQRASPLTSTI 654
Db 601 PSGVKPDLISYMPIWKFPDEEGACQPCPINCTHSCVDLDDKGCAPAEQRASPLTSTI 654

```

Search completed: January 3, 2006, 11:18:33

Job time : 55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 3, 2006, 11:16:48 ; Search time 13 Seconds

(without alignments)  
722.957 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 MELALCRWGLLALPPGA.....TFKGTPTAENPEYLGLDVPV 1255

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 57103 seqs, 7488799 residues

Total number of hits satisfying chosen parameters: 57103

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%  
Listing first 45 summaries

Database :

Published Applications\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Printed. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	US-11-022-562-213	Sequence 213, App1
2	6806	99.9	1255	US-10-770-726-62	Sequence 62, App1
3	6806	99.9	1255	US-11-113-202-10	Sequence 10, App1
4	3167	46.5	1210	US-11-113-202-6	Sequence 6, App1
5	3167	46.5	1210	US-11-145-566-1	Sequence 1, App1
6	3004.5	44.1	1308	US-11-113-202-16	Sequence 16, App1
7	2545.5	37.4	943	US-11-113-202-8	Sequence 8, App1
8	2441.5	35.8	1342	US-10-770-726-63	Sequence 63, App1
9	2441.5	35.8	1342	US-11-113-202-12	Sequence 12, App1
10	2441.5	35.8	1342	US-11-113-202-14	Sequence 14, App1
11	1874.5	27.5	419	US-11-113-202-23	Sequence 23, App1
12	1873.5	27.5	419	US-11-113-202-4	Sequence 4, App1
13	1860.5	27.3	419	US-11-113-202-2	Sequence 2, App1
14	1056.5	15.5	534	US-11-077-386-25	Sequence 25, App1
15	798	11.7	400	US-11-077-386-26	Sequence 26, App1
16	734	10.8	346	US-11-077-386-24	Sequence 24, App1
17	651	9.6	1367	US-10-995-561-538	Sequence 538, App1
18	651	9.6	1367	US-11-145-202-18	Sequence 18, App1
19	650.5	9.5	1368	US-10-995-561-539	Sequence 539, App1
20	607	8.9	987	US-10-949-720-395	Sequence 395, App1
21	606	8.9	990	US-10-821-234-1201	Sequence 1201, App1
22	584.5	8.6	976	US-10-966-483-2	Sequence 2, App1
23	583.5	8.6	1035	US-10-966-483-20	Sequence 20, App1
24	583.5	8.6	1035	US-11-021-441-4	Sequence 4, App1
25	582.5	8.5	984	US-11-113-424-60	Sequence 60, App1

26	560	8.2	456	6	US-10-966-483-34	Sequence 34, App1
27	560	8.2	456	7	US-11-021-441-18	Sequence 18, App1
28	552.5	8.1	985	7	US-11-113-424-61	Sequence 61, App1
29	551	8.1	998	6	US-10-510-524-1	Sequence 1, App1
30	543.5	8.0	987	6	US-10-770-726-61	Sequence 61, App1
31	542	8.0	879	6	US-10-770-726-78	Sequence 78, App1
32	542	8.0	995	7	US-11-113-424-62	Sequence 62, App1
33	528	7.7	479	6	US-10-966-483-36	Sequence 36, App1
34	528	7.7	479	7	US-11-021-441-20	Sequence 20, App1
35	528	7.7	490	6	US-10-966-483-42	Sequence 42, App1
36	528	7.7	490	7	US-11-021-441-26	Sequence 26, App1
37	528	7.7	497	6	US-10-966-483-38	Sequence 38, App1
38	528	7.7	497	6	US-10-966-483-40	Sequence 40, App1
39	528	7.7	497	7	US-11-021-441-22	Sequence 22, App1
40	528	7.7	497	7	US-11-021-441-24	Sequence 24, App1
41	526	7.7	983	7	US-11-113-424-59	Sequence 59, App1
42	520	7.6	1005	7	US-11-113-424-63	Sequence 63, App1
43	505	7.4	293	7	US-11-092-168-11	Sequence 11, App1
44	493.5	7.2	1035	7	US-11-113-424-20	Sequence 20, App1
45	492	7.2	1052	7	US-11-113-424-22	Sequence 22, App1

#### ALIGNMENTS

RESULT 1						
US-11-022-562-213						
Sequence 213, Application US/11022562						
Publication No. US20050249742A1						
GENERAL INFORMATION:						
APPLICANT: Shuang, Ruth M.						
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING						
FILE REFERENCE: DEN-022CN						
CURRENT FILING DATE: 2004-12-22						
PCT APPLICATION NUMBER: US/11/022,562						
PCT FILING DATE: 2004-12-22						
PCT APPLICATION NUMBER: PCT/US03/20322						
PCT FILING DATE: 2003-06-27						
PCT APPLICATION NUMBER: 60/392718						
PCT FILING DATE: 2002-06-27						
NUMBER OF SEQ ID NOS: 340						
SOFTWARE: FastSeq for Windows Version 4.0						
SEQ ID NO 213						
LENGTH: 1255						
TYPE: PRT						
ORGANISM: Homo sapiens						
US-11-022-562-213						
Query Match						
Best Local Similarity 100.0%; Score 6815; DB 7; Length 1255;						
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	MELALCRWGLLALPPGA	STGYCTGDKRLPASPETHLDMRLHYGCGVVGNTL	60		
DB	1	MELALCRWGLLALPPGA	STGYCTGDKRLPASPETHLDMRLHYGCGVVGNTL	60		
QY	61	ELTYLPTVASLSPLODIO	EVGYVLIANQROVPLQRLIVRGQLFEDNYALAVLNG	120		
DB	61	ELTYLPTVASLSPLODIO	EVGYVLIANQROVPLQRLIVRGQLFEDNYALAVLNG	120		
QY	121	DLPLNTTVPVTSAPGRL	ELRLSTELTKGSLVLIQRPOLCYOPTIMKDIFFKNNOLA	180		
DB	121	DLPLNTTVPVTSAPGRL	ELRLSTELTKGSLVLIQRPOLCYOPTIMKDIFFKNNOLA	180		
QY	181	LTLIDTNRSPACHPSPCK	SRGSCWSSSDQSLTRTVACGACRCKGPPPTDCBQC	240		
DB	181	LTLIDTNRSPACHPSPCK	SRGSCWSSSDQSLTRTVACGACRCKGPPPTDCBQC	240		
QY	241	AAAGTGPAGSCLACLP	HNHSGICELHCPALVYNTDTFESMPNPEGRYTGASCTVACP	300		
DB	241	AAAGTGPAGSCLACLP	HNHSGICELHCPALVYNTDTFESMPNPEGRYTGASCTVACP	300		

```
QY 301 YNYLSTDVSGCTLVCPHNOEVTADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
D 301 YNYLSTDVSGCTLVCPHNOEVTADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
QY 361 IOEFAGCKKITGSLAFLPESFDGPASNTAPLOPQOLOVFETLEITGYLYISAMPDLP 420
D 361 IOEFAGCKKITGSLAFLPESFDGPASNTAPLOPQOLOVFETLEITGYLYISAMPDLP 420
QY 421 DLSVFQNLQVIRGRILHNHNGAVSLTLQGLISMLGRSLRELSGSLALHHNTHLCFYVTV 480
D 421 DLSVFQNLQVIRGRILHNHNGAVSLTLQGLISMLGRSLRELSGSLALHHNTHLCFYVTV 480
QY 481 PMDQLFRNPQALHTANRPEDECVBGLACHOLCARGHCWGPPTQCVCNCSQFLRGQEC 540
D 481 PMDQLFRNPQALHTANRPEDECVBGLACHOLCARGHCWGPPTQCVCNCSQFLRGQEC 540
QY 541 VEECRVLOGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCYACAHYKDPPECVARC 600
D 541 VEECRVLOGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCYACAHYKDPPECVARC 600
QY 601 PSQVPRDLSTYMPIMKFPDEBEGACOPCINCHSCVDLDDKGCPCAPQASPLTISIAYVG 660
D 601 PSQVPRDLSTYMPIMKFPDEBEGACOPCINCHSCVDLDDKGCPCAPQASPLTISIAYVG 660
QY 661 ILVVVGLGVGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
D 661 ILVVVGLGVGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
QY 721 RKVKVLGSGAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKANKEIIDDAYVMAGVSP 780
D 721 RKVKVLGSGAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKANKEIIDDAYVMAGVSP 780
QY 781 YVSRLLGICLTSTVOLVQLMPYGCGLDHWENRGRLSODLLNMCQIAGKMSYLEBYR 840
D 781 YVSRLLGICLTSTVOLVQLMPYGCGLDHWENRGRLSODLLNMCQIAGKMSYLEBYR 840
QY 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETEHYADGGKVPDKMMALLESILRRFT 900
D 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETEHYADGGKVPDKMMALLESILRRFT 900
QY 901 HOSDWASGVYVWELMTGAKYDGI PARETIDLLLEKGBLPQPICTIIDVYMIWVKCM 960
D 901 HOSDWASGVYVWELMTGAKYDGI PARETIDLLLEKGBLPQPICTIIDVYMIWVKCM 960
QY 961 IDSECRPRRELVSFESMARDPORFVYI QNEDLPASPLDSTFYRSLLEDDMDGLVDA 1020
D 961 IDSECRPRRELVSFESMARDPORFVYI QNEDLPASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 EBYLVPOQGFPCPDAPAGAGVWHRHRSSTRSGGDLTLGLEPSEBEPASPLAPSEG 1080
D 1021 EBYLVPOQGFPCPDAPAGAGVWHRHRSSTRSGGDLTLGLEPSEBEPASPLAPSEG 1080
QY 1081 AGSDVFDGDLGMAKGLQSLPTHPDPLQYSEDPYVPLPSETTGYVAPLTCSEQEYV 1140
D 1081 AGSDVFDGDLGMAKGLQSLPTHPDPLQYSEDPYVPLPSETTGYVAPLTCSEQEYV 1140
QY 1141 NOPDVRPOPSPREBGLPAARPAAGATLERPKTSLSGKGVVADVAFCGAVENPEYTLPO 1200
D 1141 NOPDVRPOPSPREBGLPAARPAAGATLERPKTSLSGKGVVADVAFCGAVENPEYTLPO 1200
QY 1201 GGAAPQPPAPAFSPAFDNLVYWDODPEBGAPESTFKGTPAENPEYTLGLDVPV 1255
D 1201 GGAAPQPPAPAFSPAFDNLVYWDODPEBGAPESTFKGTPAENPEYTLGLDVPV 1255
```

RESULT 2  
US-10-770-726-62  
; Sequence 62, Application US/10770726  
; Publication No. US20050266409A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Brown, Eugene  
; APPLICANT: Liu, Wei

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
; TITLE OF INVENTION: CANCERS  
; FILE REFERENCE: AM101079 (031896-010000)  
; CURRENT APPLICATION NUMBER: US/10/770,726  
; CURRENT FILING DATE: 2004-02-04  
; NUMBER OF SEQ ID NOS: 48640  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 62  
; LENGTH: 1255  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-10-770-726-62
```

Query Match 99.9%; Score 6806; DB 6; Length 1255;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MELAALCRWGLLLALPPGAASSTOVCTGDMKRLPASPEHLDMLRLYGGCQVQGNL 60
D 1 MELAALCRWGLLLALPPGAASSTOVCTGDMKRLPASPEHLDMLRLYGGCQVQGNL 60
QY 61 ELTYLPTNASLSPLODIQEVQGYVLIANOVROYPLORLIRVGTQLPEDNYALAVLDNG 120
D 61 ELTYLPTNASLSPLODIQEVQGYVLIANOVROYPLORLIRVGTQLPEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKSGVLIQRNPOLCYQDTIIMKDI FHKNNQLA 180
D 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKSGVLIQRNPOLCYQDTIIMKDI FHKNNQLA 180
QY 181 LTLIDTNRSAKCHCSMPCKSKRCWGBSSBECQSLTRVNCAGCGARCKGRLPTDCHEQC 240
D 181 LTLIDTNRSAKCHCSMPCKSKRCWGBSSBECQSLTRVNCAGCGARCKGRLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLCLPHNHSICELHCPALVYNTDTPESMPNBEGRYTFGASCVTACP 300
D 241 AAGCTGPKHSDCLCLPHNHSICELHCPALVYNTDTPESMPNBEGRYTFGASCVTACP 300
QY 301 YNYLSTDVSGCTLVCPHNOEVTADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
D 301 YNYLSTDVSGCTLVCPHNOEVTADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
QY 361 IOEFAGCKKITGSLAFLPESFDGPASNTAPLOPQOLOVFETLEITGYLYISAMPDLP 420
D 361 IOEFAGCKKITGSLAFLPESFDGPASNTAPLOPQOLOVFETLEITGYLYISAMPDLP 420
QY 421 DLSVFQNLQVIRGRILHNHNGAVSLTLQGLISMLGRSLRELSGSLALHHNTHLCFYVTV 480
D 421 DLSVFQNLQVIRGRILHNHNGAVSLTLQGLISMLGRSLRELSGSLALHHNTHLCFYVTV 480
QY 481 PMDQLFRNPQALHTANRPEDECVBGLACHOLCARGHCWGPPTQCVCNCSQFLRGQEC 540
D 481 PMDQLFRNPQALHTANRPEDECVBGLACHOLCARGHCWGPPTQCVCNCSQFLRGQEC 540
QY 541 VEECRVLOGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCYACAHYKDPPECVARC 600
D 541 VEECRVLOGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCYACAHYKDPPECVARC 600
QY 601 PSQVPRDLSTYMPIMKFPDEBEGACOPCINCHSCVDLDDKGCPCAPQASPLTISIAYVG 660
D 601 PSQVPRDLSTYMPIMKFPDEBEGACOPCINCHSCVDLDDKGCPCAPQASPLTISIAYVG 660
QY 661 ILVVVGLGVGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
D 661 ILVVVGLGVGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
QY 721 RKVKVLGSGAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKANKEIIDDAYVMAGVSP 780
D 721 RKVKVLGSGAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKANKEIIDDAYVMAGVSP 780
QY 781 YVSRLLGICLTSTVOLVQLMPYGCGLDHWENRGRLSODLLNMCQIAGKMSYLEBYR 840
D 781 YVSRLLGICLTSTVOLVQLMPYGCGLDHWENRGRLSODLLNMCQIAGKMSYLEBYR 840
```

QY 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLDIDETEHADGKVPDKMMALSSILRRFT 900  
DB 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLDIDETEHADGKVPDKMMALSSILRRFT 900  
QY 901 HOSDWSYGVTVWELMTFGAKPYDGI PARBEPDLLEKGBRLPQPICTIDVYMIWKKCM 960  
DB 901 HOSDWSYGVTVWELMTFGAKPYDGI PARBEPDLLEKGBRLPQPICTIDVYMIWKKCM 960  
QY 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGAPSLDSTFYRSLLDDMDGLVDA 1020  
DB 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGAPSLDSTFYRSLLDDMDGLVDA 1020  
QY 1021 BEYLVPOQGFPCPDPAFGAGVWVHRHRSSTRSGGDLTLGLBSEBEPAPSLAPSEB 1080  
DB 1021 BEYLVPOQGFPCPDPAFGAGVWVHRHRSSTRSGGDLTLGLBSEBEPAPSLAPSEB 1080  
QY 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEBPTVPLPSETDGYVAPLTCSPQEVY 1140  
DB 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEBPTVPLPSETDGYVAPLTCSPQEVY 1140  
QY 1141 NOPDVRPQPSREBGLPAARAGATLBRKTLSPGKGVVDVAFGAVENPEYLTPO 1200  
DB 1141 NOPDVRPQPSREBGLPAARAGATLBRKTLSPGKGVVDVAFGAVENPEYLTPO 1200  
QY 1201 GGAAPQHPHPAFSFAFDNLYYMDQDPPERGAPSTFKGTPTAENPEYLGIDVPV 1255  
DB 1201 GGAAPQHPHPAFSFAFDNLYYMDQDPPERGAPSTFKGTPTAENPEYLGIDVPV 1255

## RESULT 3

US-11-113-202-10  
; Sequence 10, Application US/11113202  
; Publication No. US20050272637A1  
; GENERAL INFORMATION:  
; APPLICANT: Clinton, Gail M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1  
; FILE REFERENCE: 49321-136  
; CURRENT APPLICATION NUMBER: US/11/113,202  
; PRIOR FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/590,473  
; PRIOR FILING DATE: 2004-07-23  
; PRIOR APPLICATION NUMBER: US 60/564,893  
; PRIOR FILING DATE: 2004-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 10  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-113-202-10

Query Match 99.9%; Score 6806; DB 7; Length 1255;

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MELAALCGWGLIALLPQGAASVCTGTDMKRLRPASBETHLDMRLHYGCCVYQGANL 60  
DB 1 MELAALCGWGLIALLPQGAASVCTGTDMKRLRPASBETHLDMRLHYGCCVYQGANL 60  
QY 61 ELTYLPTNASLFLDDIOEVGVYLIAHNOVQVLOLRIRYRGQLPEPDNALAVLNG 120  
DB 61 ELTYLPTNASLFLDDIOEVGVYLIAHNOVQVLOLRIRYRGQLPEPDNALAVLNG 120  
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLLIQRNPOLCYODTILMKDIFHNQOLA 180  
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLLIQRNPOLCYODTILMKDIFHNQOLA 180  
QY 181 LTLITNNSRACHPCSPMKSGRCWGESSEDCOSILTRTYCAGGACRCAGPLPTDCCHEOC 240  
DB 181 LTLITNNSRACHPCSPMKSGRCWGESSEDCOSILTRTYCAGGACRCAGPLPTDCCHEOC 240

QY 241 AAGCTGPKHSDCLACIAHFNHSGICELACPALVTYNTDTPESMNPBEGRTYFGASCVTACP 300  
DB 241 AAGCTGPKHSDCLACIAHFNHSGICELACPALVTYNTDTPESMNPBEGRTYFGASCVTACP 300  
QY 301 YNYISTDVSGCTLVCPPLNQEVTAEEDGTORCEKSKPCARVYCYGIMHLEBRVAVTSAN 360  
DB 301 YNYISTDVSGCTLVCPPLNQEVTAEEDGTORCEKSKPCARVYCYGIMHLEBRVAVTSAN 360  
QY 361 IOEFAGCKKIKESIALFLPESFPDGPASNTAPLOPQLOVFEETLEBITGLYLSAMPDSLAP 420  
DB 361 IOEFAGCKKIKESIALFLPESFPDGPASNTAPLOPQLOVFEETLEBITGLYLSAMPDSLAP 420  
QY 421 DLVSFQNLQVIRGRILIHNGAYSILTQGLGISWLGRLSRELSGSLALIHNTHLFCFVTV 480  
DB 421 DLVSFQNLQVIRGRILIHNGAYSILTQGLGISWLGRLSRELSGSLALIHNTHLFCFVTV 480  
QY 481 PWDQLFRPHQALLHTANRPEDECVSGIACHQLCARGHCWGPPTQCVNCSQFLRGQBC 540  
DB 481 PWDQLFRPHQALLHTANRPEDECVSGIACHQLCARGHCWGPPTQCVNCSQFLRGQBC 540  
QY 541 VEECRVLOGPREYVNAHCLPCHPECPQNGSVTCFGEADQCVACHYKDPFCVAVAC 600  
DB 541 VEECRVLOGPREYVNAHCLPCHPECPQNGSVTCFGEADQCVACHYKDPFCVAVAC 600  
QY 601 PSGVPRDLSTYMPITWKFPEBEGACQPCPINCSTHSCVDLDDKGCAPQORASPLTSIISAVVG 660  
DB 601 PSGVPRDLSTYMPITWKFPEBEGACQPCPINCSTHSCVDLDDKGCAPQORASPLTSIISAVVG 660  
QY 661 ILVVVIGVFGIILIKRQOKIRKTYMRLLQETELVEBLTPSGAMPNOQMRILKETEL 720  
DB 661 ILVVVIGVFGIILIKRQOKIRKTYMRLLQETELVEBLTPSGAMPNOQMRILKETEL 720  
QY 721 RKVKVLGSGARCTYVKGIMIPDGENVKI PVAILKVLRENTSPANKETIIDEAYVMGVGSP 780  
DB 721 RKVKVLGSGARCTYVKGIMIPDGENVKI PVAILKVLRENTSPANKETIIDEAYVMGVGSP 780  
QY 781 YVSRLLGICLNSTVQLVQLMPYGCILDHVENRGRLSQDLIANNQOIAKMSYLEBVR 840  
DB 781 YVSRLLGICLNSTVQLVQLMPYGCILDHVENRGRLSQDLIANNQOIAKMSYLEBVR 840  
QY 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLDIDETEHADGKVPDKMMALSSILRRFT 900  
DB 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLDIDETEHADGKVPDKMMALSSILRRFT 900  
QY 901 HOSDWSYGVTVWELMTFGAKPYDGI PARBEPDLLEKGBRLPQPICTIDVYMIWKKCM 960  
DB 901 HOSDWSYGVTVWELMTFGAKPYDGI PARBEPDLLEKGBRLPQPICTIDVYMIWKKCM 960  
QY 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGAPSLDSTFYRSLLDDMDGLVDA 1020  
DB 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGAPSLDSTFYRSLLDDMDGLVDA 1020  
QY 1021 BEYLVPOQGFPCPDPAFGAGVWVHRHRSSTRSGGDLTLGLBSEBEPAPSLAPSEB 1080  
DB 1021 BEYLVPOQGFPCPDPAFGAGVWVHRHRSSTRSGGDLTLGLBSEBEPAPSLAPSEB 1080  
QY 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEBPTVPLPSETDGYVAPLTCSPQEVY 1140  
DB 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEBPTVPLPSETDGYVAPLTCSPQEVY 1140  
QY 1141 NOPDVRPQPSREBGLPAARAGATLBRKTLSPGKGVVDVAFGAVENPEYLTPO 1200  
DB 1141 NOPDVRPQPSREBGLPAARAGATLBRKTLSPGKGVVDVAFGAVENPEYLTPO 1200  
QY 1201 GGAAPQHPHPAFSFAFDNLYYMDQDPPERGAPSTFKGTPTAENPEYLGIDVPV 1255  
DB 1201 GGAAPQHPHPAFSFAFDNLYYMDQDPPERGAPSTFKGTPTAENPEYLGIDVPV 1255

## RESULT 4

US-11-113-202-6  
; Sequence 6, Application US/11113202  
; Publication No. US20050272637A1



```

Db 74 YDLSFKTIQEVAGVYLIALNTERIPLLENIQIRGNMYENSVALAVLSND----- 126
Qy 129 VTGASPGGLRELQRLSLTEILKGVLIQBNPOLCYODTILMKDIFHKNNQALTLIDTR 188
Db 127 ---ANKTELKEIPMANLOEILHGAVRFSNNPALCVESIQMDIYSSDPLSMKMDPQNH 183
Qy 189 SRACHPCSPMKSGHSCWGESSEDCSLTRTVCAAGCA-RCKGPLPTDCCHQCAAGCTGP 247
Db 184 LGSCKCKDPSCPNCSGWSGABEENCKQLTKILCAQCSGCRKSGSPDCCHNCAAGCTGP 243
Qy 248 KHSDDLACHFNHSGICELHCPALVTYNTDTFESMPNBEGRTRTPGASCTYACPYNYLSTD 307
Db 244 RESDCLVRKRDEATCKDTCPPMLYNYPTTYQMDNPEGKXSFATCTKCKPRYVYVD 303
Qy 308 VGSCTLVCPRLHNOETAEEDGTORCEKSPKARVCYGLMEHLREVRATVANSANIOEFGAC 367
Db 304 HGSCTRACGADSYEN-EEEDGVAKCKKCBGPKVCKNGIGIGFCKDLSINAINIHFKNC 362
Qy 368 KKIIFGSLAFLEPSFDGDPASNTAPLQBPOLQVFETLEITGYLYISAMPDSLPLDSVFON 427
Db 363 TSIISGDHLILPVAFRGDSFTHTPPLDPOGLDILKTVEKITGFLLIQAMPENRTDLHAFEN 422
Qy 428 LQVIRGRILHNGAYSLTLOGLISWLGRLSLRELSGLALIHNTHLCPVHTVPDQLR 487
Db 423 LEIRGRTRKHQGFSLAVVSLNITSLGRSLKEISDGVIIISGNKLCYANTINMKKLF 482
Qy 488 NPHOALHTANRPEDECVGEGGLACHQLCARHSCWGPPTOCVNSQFLRGCEVCECHYL 547
Db 483 TSGQCTKILISNNGENSCATGOVCHALCSPEGCWGPBPDVCSKNVSGRGECVCKNUL 542
Qy 548 QGLPREYVNAHCLPCHPECOFONGSVTCFGEADQCAVACHKDPFCVACSPGVKPD 607
Db 543 EGBREFEVENSBCIQCHPECLPQAMNITCTGGRPNCTQCAHYIDGPHCVKTCRPAWGME 602
Qy 608 LSYWIKFPRBEAGACQPCPINCHSCVDLDDKCPBAPORASPLTISIISAVG---ILIV 664
Db 603 NNLT-VMKYADAGHCHLCHPCTGYCTGPGLEGCPYMGPKP--SIATGMGALLLILV 659
Qy 665 VILGVVPGILIKRROKIRKYTMRLLOETELVEPLTSGAMPNOAKRILKETELRYK 724
Db 660 VALGCG---LPMRRNHIVKRLRLLORELVEPLTSGEAPNALILKLETFPKIK 716
Qy 725 VLGSAGFTVYKGIWIPGENVYKIPVAILKVLRENTSPYANKELIDEAVYMAVGSPVYSR 784
Db 717 VLGSAGFTVYKGIWIPGENVYKIPVAILKVLRENTSPYANKELIDEAVYMAVSPVYSR 784
Qy 785 LIGICLSTVOLVLOLMPYGCILDHVRENGRLSGODILNMCQIAKMSYLEDVRLVYR 844
Db 777 LIGICLSTVOLVLOLMPYGCILDHVRENGRLSGODILNMCQIAKMSYLEDVRLVYR 844
Qy 845 DLARBNVLKSPNHVKTIDPGIARLDIDETEVHADGGKVPYKMMALSSILRRFTHSD 904
Db 837 DLARBNVLKSPNHVKTIDPGIARLDIDETEVHADGGKVPYKMMALSSILRRFTHSD 904
Qy 905 VMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRILPOPICITIDVYMIWKCMIMISE 964
Db 897 VMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRILPOPICITIDVYMIWKCMIMISE 964
Qy 965 CRPRFRELVSFRRMARDPQRFVVIQ-NEDUGPASPLDSTFYRSLLEDNDKMDLVDAEY 1023
Db 957 SRPKRELILIEFSKARDPORLYVIGDBRMHLPSPTSNFRAIMDEEDMDVDVDAEY 1016
Qy 1024 LVPOGFGFCPRPARAGGMVHHRRSSSTRSGGDLTGLBPSSEEARSPAPSEGGS 1083
Db 1017 LVPOGFGFCPRPARAGGMVHHRRSSSTRSGGDLTGLBPSSEEARSPAPSEGGS 1083
Qy 1084 DVFPDDLGMGAAGLQSLPTHDPSPLOQRYSEDPVPLPSET--DGYVAPLTCSPQPEYVN 1141
Db 1043 N-INSTVACIRANGIQCSPRIKEDSFLQRYSSBPTCALBEDSIDDFL-----PPEYIN 1094
Qy 1142 QPDVAPPPSPREGLPAPAPAGATL-----ERPPTLSPGKNGVVKDVPFAGAVENDEY 1196
Db 1095 Q-----SVP-KRPRAGSVQNPVYHNPQPLNPAAPSRDPHQD--PHSTAVGNPEY 1138

```

```

Qy 1197 L-TPGGAAPQHPRPAPSPADNLXYWD-----DP-----PERGAPSTFKG 1239
Db 1139 LNTVQ-----PTCVNSTFSDPAPHAQKSGHQSISLNDPVOQDFFPKKAPNGIFKG 1189
Qy 1240 TPTAPNPEYL 1249
Db 1190 S-TAPNAEYL 1198

RESULT 6
US-11-113-202-16
; Sequence 16, Application US/11113202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; FILE REFERENCE: 49321-136
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US/11/113,202
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; PRIOR FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16
; LENGTH: 1308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-16

Query Match 44.1%; Score 3004.5; DB 7; Length 1308;
Best Local Similarity 45.6%; Pred. No. 2.2e-183;
Matches 614; Conservative 183; Mismatches 375; Indels 175; Gaps 27;

9 MGLLALLPFGAA-----STOYCTGDMKRLRASPETHLDMRLHLYOGCOVVOGNTLELY 64
8 WWSVSLVAAGTVQPSDSQSVACAGTENKLSLSLEQOYRALRYKTYEMCEVVMGHEITS 67
65 LPTNASLFLDDIOEVQGVLIHANVOVPLQRLIRVGTQLPEFNYALAVLDNGDPLN 124
68 IEHRNDSFLHSVREVTGYVALNQFRYLPLENRIIRGKLVYDRVALAIFLYNRKDG 127
125 NTPYTGASPGGLRELQRLSLTEILKGVLIQBNPOLCYODTILMKDIFHKNNQALTLI 184
128 NF-----GLOELGIKNLTELINGGVYDQNFCLCYADTINHODIVRNPWPSNLTLY 178
185 DTNRSRACHPCSPMKSGHSCWGESSEDCSLTRTVCAAGC-ARCKGPLPTDCCHQCAAG 243
179 STNSSGCGRCHKSGTG-RCKGPEFNHCQTLIRTVCAQCGRCYGPVYSDCHRECKAG 237
244 CTGPRGSDCLALCFHNHSGICELHCPALVTYNTDTFESMPNBEGRTRTPGASCTYACPYNY 303
238 CSGPRDTCFACQAMFNPNDSGACVTCQPFQFVYNNPTTFQLEHFNPAKYTGACVCKCPHNF 297
304 LSTDVSGCTLVCPRLHNOETAEEDGTORCEKSPKARVCYGLMEHLREVRATVANSANIOE 363
298 V-VDSSSCVPRACPSGSKMEV-BENGIKMKCKPCTDICPKKACDIDGTGSLMSADTVVSSNIDK 355
364 FAGCKKIFGSLAFLEPSFDGDPASNTAPLQBPOLQVFETLEITGYLYISAMPDSLPLDS 423
356 FINCKTNGNLIFLVGTGHDPPYNAIEAIDCKLVNFTVEITGFLIIGSMPPRMTDPS 415
424 VFQNLQVIRGRILHNGAYSLTLOGLISWLGRLSLRELSGLALIHNTHLCPVHTVPD 483
416 VFSNLVTGIGVLSGSLILTKQGGITSLOFQSKELISAINIYITDMSNICYHTIWT 475
484 QLRNPHOALHTANRPEDECVGEGGLACHQLCARHSCWGPPTOCVNSQFLRGCEVCEY 543
476 TLFPSTINQRIIVIRNKRKAENCTAGMVCNHLSSSDGCGWGPBPDCLSCRPSRGRICIES 535

```



```

Qy 544 CRYLQGLPREYVNAHCLPCHEPCOP-ONGSVTCGPEADOCVACAHKDPFCVAPRPS 602
Db 536 CNLDGGEREPENGSI CVCEDPCCEKEDGLTCHGPRDNTCKSHKQDGNVCYKECPD 595
Qy 603 GVKPDLSTYMPIWKFPDEEGACOPCPINCTHSCVDLDDKGC-----PAEQRASPL 651
Db 596 GLQGANSP--IFKXADPDRECHPCHPNCTGCGNCPHTSHDCIYYPWTHGSTLPQHAR----- 649
Qy 652 TSIISAVV--GILLVVLGVVFGILIKRQCKIRKTYMRRLLOETELVEPLTPSGAMPQ 709
Db 650 TPLAAGVIGGLFVLIVGLTFVAVYRRKSKIK-KRRALRRFL-ETELVEPLTPSGTANQ 707
Qy 710 AQMRLLKTELKRVKVLGSGAFGYKGIWIPDGENVKIPVAIKYLRNTPSKANKETLD 769
Db 708 AQRLKTELKELKRVKVLGSGAFGYKGIWIPDEGTVAIPVAIKYLRNTPSKANKETLD 767
Qy 770 EAYVAGVSPVSRRLGICLTSTVQVLTOLMPYGCCLDHYENRGRLSODLNMCMQI 829
Db 768 EALIMASMDHPLVRLVLGVCSPITQVLTQMPHGCCLLEYHEKDNIGSOLLNMVCQI 827
Qy 830 AKGMSYLEDVRLVARDLAARNVLKSPNNVKITDGLARLDDIDETVHADGKVPITKM 889
Db 828 AKGMVYLERRLVHRDLAARNVLKSPNNVKITDGLARLLEGEKEXNADGKMPIKM 887
Qy 890 ALESLRRRPTHQSDVMSGYVTWELMTFGAKPYDGIPIAREIPDLLEKGERLPOPICTI 949
Db 888 ALECHYKFKTHQSDVMSGYVTWELMTFGKPYDGIPIAREIPDLLEKGERLPOPICTI 947
Qy 950 DVMYIMVCMWIDSECRPRFRELVEFSRMAPDQRFVYIIONED-LGPAASPLDSTFYSL 1008
Db 948 DVMYIMVCMWIDSECRPRFRELVEFSRMAPDQRFVYIIONED-LGPAASPLDSTFYSL 1007
Qy 1009 LEDDMGLVDAEELVLPQCGFCDDPAAGAGMHHHRSSSTSGGDLTLGLEPBE 1068
Db 1008 LEDDELDMDMAEELVLP-QAFNIPPP-----IYTSARLDSNMS-----EIGHSPPA 1055
Qy 1069 EAPRS-----PLAP--SEGAGSVFPGDLGMAKAGLQ 1099
Db 1066 YTPMSGNQFVRDGFAPAEQGVSVYRAFTSTIPAPVAQGTAFIPDSCNGTLRKPV 1115
Qy 1100 SLPTHPSPRLQRYSEDPTVPLPS-----ETDGYVAPLTCSPQPEYVNOQDVRQPPSP 1152
Db 1116 APHVQEDSTQYSADPTVPAERSPRGELDEBGMTRMDRKPQGEYLNPAV----- 1167
Qy 1153 REGPLPAARPAAGATLERPKTSLSPGKGVYKDVAFAGAVENPEYLTPOGGAAPQHPPPA 1212
Db 1168 -ENPFVSR-----KNGDLQ-----ALDNEPEYNASNG-----PPKA 1198
Qy 1213 -----FSPAFDNLYYMDQDPPERGA--PPST 1236
Db 1199 EDEYVNEPLYLNTFANTLGKAEYLNKNNLISMEPKAKKAFNDPNWNSLPPRSTLQHPDY 1258
Qy 1237 FKGTPT-----AENPEYL 1249
Db 1259 LQEVSTKVFYKQNGRIPIVAENPEYL 1285

RESULT 7
US-11-113-202-8
; Sequence 8, Application US/11113202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; APPLICANT: Shantleh, Lara
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; TITLE OF INVENTION: RECEPTOR AND ERBB RECEPTORS
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893

```

```

; PRIOR FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8
; LENGTH: 943
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-8

Query Match      37.4%; Score 2545.5; DB 7; Length 943;
Beet Local Similarity 51.3%; Pred. No. 2e-154;
Matches 510; Conservative 136; Mismatches 245; Indels 103; Gaps 19;

Qy 284 NPEGRYFGASCVYACACPNYLSITDVGSCITVCPRLHNOBVTAEADGTORCEKSKPCARVCY 343
Db 13 NPEGRYFGANCVYKCCRNLYVTVDHSGCVACGADSYEM-EDGVKCKCKECCGCRKCN 71
Qy 344 GLGMEHLREVAVTSANIQEPAQCKTIGSLAFIPSEFDGDPASNTAPLOEQVETL 403
Db 72 GIGIGEFKDSISINATNIKFKNCTSIGDLHILPVAFRGDSFHTPPLDPOEIDLTKTV 131
Qy 404 EETIGYLYISAMPDLSLDFQNLQVIRGRILHNGAVSLTQGLGISLGLRSLRGLS 463
Db 132 KEITGFLIQMPENRITDLHAFENLEIRGKTKQHQPSLAVSLNLTSLRSLKEISD 191
Qy 464 GLALIHNTLCLFHTVPMQDLFRNPHQALHTANRDEDECVGBGLACHOLCARGHCMWP 523
Db 192 GDVYISGNKNCVYANTINMKLFTSGQKTKIISNRGENSKATGVCYCHALCPEGCGWP 251
Qy 524 GPTCCVCSQFLRQCEVCEBQVLOGLPREYVNAHCLPCHEPCOPQNGSVTCGPEADQ 583
Db 252 EPRDCVCSRNYSRREGCVCKNLLEGPRFEVENSECICQCHBCLPQAMNITTCGRGPDN 311
Qy 584 CVAQAHKDPFCVACRPSGVKPDLSYMPIWKFPDEEGACQPCPINTCHSCVDLDDKGC 643
Db 312 CIQCAHYTDGPHCVKTCGPAGWGENNTL-VKVIADAGHVCHLCPNCTCTGCTGELBGP 370
Qy 644 AEQASPLTSIISAAG---ILLVVLGVVFGILIKRQCKIRKTYMRRLLOETELVEPL 700
Db 371 TNGRKIP--STATMGVALLLLVVALGIG--LPMRRRIHVRKTRRLRRLQERLVEPL 425
Qy 701 TPGAMPNOAQMLKETELKRVKVLGSGAFGYKGIWIPDGENVKIPVAIKYLRNTPS 760
Db 426 TPSEAPPAQALRIKETEFEKIKVYKVLGSGAFGYKGIWIPDEGKVPVAIKYLRNTPS 485
Qy 761 PKAKEILDEAYVNAAGVSPVSRRLGICLTSTVQVLTOLMPYGCCLDHYENRGRLSQ 820
Db 486 PKAKEILDEAYVNAAGVSPVSRRLGICLTSTVQVLTOLMPYGCCLDHYENRGRLSQ 545
Qy 821 DLNMCMQIAKMSYLEDVRLVHRDLAARNVLKSPNNVKITDGLARLDDIDETVHAD 880
Db 546 YLNMCMQIAKMSYLEDRVLVHRDLAARNVLVTPQVVKITDGLAKLGLAEKEVYAE 605
Qy 881 GKGVPKIMMALESILRRRPTHQSDVMSGYVTWELMTFGAKPYDGIPIAREIPDLLEKGER 940
Db 606 GKGVPKIMMALESILHRYTHQSDVMSGYVTWELMTFGSKPYDGIPIAREISLIEKGER 665
Qy 941 LPQPICTIDVYIMVCMWIDSECRPRFRELVEFSRMAPDQRFVYIIONEDGPAASP 999
Db 666 LPQPICTIDVYIMVCMWIDSECRPRFRELVEFSMAADPOKVLVYIQSDERHNLPS 725
Qy 1000 LDTFYSLLEDMDGLVDAEELVLPQCGFCDDPAAGAGMHHHRSSSTSGGDL 1059
Db 726 TDSNFRALMDEBMDVDVDAEVLIPQCGF----- 757
Qy 1060 TLGLEPSEEARPSPLAPSEAGSDVFDGDLGMAKAGLQSLPTHDPSPRLQRYSEDPTVP 1119
Db 758 -----SSPSTRPPLSLSLATSN--NSTVACIDRNLOSCPIKEDFRLQRYSDPFGA 809
Qy 1120 LPSET--DGYVAPLTCSPQPEYVNOQDVRQPPSPRBRGLPAARPAAGTTL-----ERPXT 1172
Db 810 LTEDSIDTFL-----FVPEYINO-----SVP-KRPAGSVQNPVYHNOPLN 849

```

Query 1173 LSPGNQGVKQVAFPAFCGAVENPEYL-TPQGGAAQPPHPAFAFPAFDLTYMDQ----- 1225  
Db 850 PAFSDPHYQD--PHSTAVGNPEYLNTVQ-----FTCVNSTFDSPAHMAQKSHQI 898  
Qy 1226 ---DP-----PERGAPSTFKGTPTANPEYL 1249  
Db 899 SLNDPDDYQDPPFKPAKNGIFPKGS-TAENPEYL 931

RESULT 8  
US-10-770-726-63  
; Sequence 63, Application US/10770726  
; Publication No. US20050266409A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Brown, Eugene  
; APPLICANT: Liu, Wei  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
; TITLE OF INVENTION: CANCERS  
; FILE REFERENCE: AM101079 (031896-010000)  
; CURRENT APPLICATION NUMBER: US/10/770.726  
; CURRENT FILING DATE: 2004-02-04  
; NUMBER OF SEQ ID NOS: 48640  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 63  
; LENGTH: 1342  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-770-726-63

Query Match 35.8%; Score 2441.5; DB 6; Length 1342;  
Beet Local Similarity 41.0%; Pred. No. 1.3e-147;  
Matches 539; Conservative 182; Mismatches 457; Indels 137; Gaps 34;

Qy 10 GLLALALPFGAA--STOVCTGTDMLRLPASPTHLDMLRLHYOGQGVVQGNLELYLPT 67  
Db 11 GLTSLARGSEVGNQAVCPGTGLNGISTGDAENQYQTLKYLRCERVGNLEIVLTGH 70  
Qy 68 NASLSPLODIOEVQGYVLIANNOVQVPLQRLRIYRGTLFEDNYALALDNGDPLNNT 127  
Db 71 NADSLPQWIEVYGVYLVANNEFTPLPLRAVVRGTVYDGKFAIFVM-----LNYNT 125  
Qy 128 PYTGAFCGGLRELQRLSTLEILKGGVLIQRNPOLCYQDTILMKQIFPKNNQALTLIDTN 187  
Db 126 ----NSSHAFLQLRLTQLTEILSGGVYIEKNDKLCMDITIMRDIIVDRD---AEIIVKD 178  
Qy 188 RSRACHPSPCKSRCKWGESEDCQSLTRIVCAGGC-ARCKGDLPTDCGECQCAAGTGT 246  
Db 179 NGRSCCPHEVCKG-KCMGPGSEDCQTLTKITCAPQCGHCFGNPNOCCCHDEACGGSG 237  
Qy 247 PKHSDCLACLHFNNSGICELHCPALVTYNTDTFESMPNBSRGYTFGASCVTACPYNTLST 306  
Db 238 PQDIDCFACRHFNDSGACVPRCPQVYVYNKLTFLQLEPRPHTKYQGVGVASCPHFV-V 296  
Qy 307 DVSGCTTLCPLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREYAVTSANIQEPAG 366  
Db 297 DOTGCVRACPDKHVD-KNGLKMCCEPCGGLCPACEGTGSG--SRQTVDSNIDFVN 353  
Qy 367 CKKTFGSLAFPESEFDGPASNTAPLQPEOLQVETLEITGYLYIAMPRLSDLVQF 426  
Db 354 CTKLGNIDFLITGLNGDPMHKIPALDEKLNVRVETIETGYLNIQSWPMMNFVSFS 413  
Qy 427 NLQVYRGRILHNGAVS-LTLOGLGISWGLSLRELSGLALLIHNTHLCFVHTVPMDQL 485  
Db 414 NLTTIGGSLVNRGSLIMKNLWTSIGPFSLEISAGRIYISANQLCHHSILNMTKV 473  
Qy 486 FRNPHQALLHTA-NRPEDECVGEGLAHQLCARGHCMWPGFTQCVNCSQFLRGQECYBEC 544  
Db 474 LRGPTEERLDIKHNRPRRDCVAEGKVCDPCLSSGGCMWPGGQCLSCMYSRGGCVYTHC 533  
Qy 545 RVLQGLPREVYNAHHCPLRPECQPPONGSVYCPGEADQCVACAHAYDPRPCVARCSPGV 604  
Db 534 NFLNGEPEEFHABECSCHPCEQPMGGTATCNDSGSTCAQCAHFRGPRGVSSCPHGV 593

```

Cy 605 KPDLSTWPIKPEPBEGACQCPINCHTSHSCYDLDKCCPRAEQR-----SPLSTIISAVVG 660
Db 594 LG--AKGYKYPDVNOCRCPEHCNCTOGCKGPELODLOGTTLVIGKTHLTMALVIAG 651
Cy 661 ILTVVLGVVFGILIKKROOKIR-KYMRRLIQETELVEPLTPSGAMPNOAMRILKETE 719
Db 652 --LVIFMMLGQFTLWYRGRIIOKRAMRRLERGESIEPLDPB-EKANKYLANIPRETE 708
Cy 720 LRKQVLSGSGFQTVYKGIWIPDGENYKIPALVLEBNTSPKANKEILDEAYMAGVGS 779
Db 709 LRKIKVLGSGVFGVHKGWIPBESGSIKIPVICIVIDKSGOSFOAVTHMLAIGSLDH 768
Cy 780 PYVRLILGICLTSTVOIYVTOLMYPGCLLDHYRENRGLSGODLLNMCQIAGKSGYEDV 839
Db 769 AHVIRLGLCGSSLQDVTQVPLGSLDHRQKRLGFPOLLNMGVQIAGKGYVIEBH 828
Cy 840 RLVRDLAARNVVLKSPBNHYKITDFGLARLIDLIDETEHADGKVPKIMWALBESILRRF 899
Db 829 GMYHRNLAAARNVLLKSPSQVQVADFGVADLLPDDKQLVSEAKTPIKMWALBESIHFGKY 888
Cy 900 THQSDVMSYGYTWMELMTFGAKPYDGIIPAEIIPULEKGERLPPPICTIDVYIMWKW 959
Db 889 THQSDVMSYGYTWMELMTFGAEPYAGRLAEVPLLEKGERLAPQICTIDVYVWVWKW 948
Cy 960 MIDCECPREFELVSEFSRMAPDQRFVYVIONEDLGA-----SPDSTFYRSLDEDDMG 1016
Db 949 MIDENIIPTEFKELANESTRMAKDPKRYLVIKRES-GEGLIAGPPEPHGLTNKGLEVELEP 1007
Cy 1017 LVDAEYLVPOQGFPCDPAPGA---GGVYHHRHRSSTRS-----GGGDL 1059
Db 1008 ELDDLDLLEAEDVLAITTLGSLSLSPVGTILNRPGSQSLSPSSGYPMNQMGISGSCQ 1067
Cy 1060 TLGISEPBEERPRPLA-----PSEGLASDVPEGLGMAKGLDS-----LPTHDP 1107
Db 1068 ESAVSGSSERCP-PSVSLHPMPRCGLASBESBEGHV-TGSEALEQEKVSMCRSRSRSPR 1125
Cy 1108 P-----LQKYS-EDPTVPV-----PSHDDGVVAPLTCSPOPEYVNOPEVRPO-PPSP 1152
Db 1126 PRGSAHSHQSHSLTIVTPLSPGLEEDVNGYV-----MDPTHKGRPPS 1172
Cy 1153 REGPLPAARPAAGATLERPKTLSPGKNGV-----KDVAFGAVENPEYVLTPOGGAPO 1206
Db 1173 REGTLS-----SVGLSSVLDTHEEDED-----BEVEYNNRRRRHSP- 1208
Cy 1207 PHPPAPSSPAFDNLYMD-----ODPPERGAPSTPFKGTAAENPEYL 1249
Db 1209 PHPRPSSLBELGYEMDVGSJLSASLGSTQSCPLHPVPIMPATGTPDEDEYEM 1263

RESULT 9
US-11-113-202-12
; Sequence 12, Application US/11113202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; TITLE OF INVENTION: RECEPTOR AND ERBB RECEPTORS
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113, 202
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590, 473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564, 893
; PRIOR FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 1342
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-12

```

Query Match 35.8%; Score 2441.5; DB 7; Length 1342;  
 Best Local Similarity 41.0%; Pred. No. 1.3e-147;  
 Matches 539; Conservative 182; Mismatches 457; Indels 137; Gaps 34;

```

10 GLLALLPPGAA--STGYCTGTMKRLPABPETHLDMRLHYGQCQVYQGNLEITYLPT 67
11 GLFLSLARGSEVGNQAVCPGTLNGLSVTGDENQYOTLYKLYERCEVVMGNLEIVLIGH 70
68 NASSLFLDIOEVQGYVLIANQVAVPLQRLRYRGQLEPENVALVLDNGPLNNTT 127
71 NADSLFLQWIREVGYVLIANNEFSTLPNLRVVRGQYVQDGKFAIFVM-----LNNNT 125
128 PVTGASPGLELQRLSLTEILKGGVLIQRPOLCYODTIIMKDI FHNKQALATLIDTN 187
126 ----NSSHALRQLRLTQLTLEILSGGYIEKNDKLCMDTIMRDIVRDRD---AEIVK 178
188 RSRACHPCSPMKSGRCWGESEDCQSLTRTVCAAGC--ARCKGRLPTDCCHGQCAAGCTG 246
179 NGRSCPCHGVCKG--RCWGPGESEDCQTLTKITICAPQCNHCGFGRPNPNCCHDECAAGCGSG 237
247 PKHSPLACLFHNSGICELHCPALVTYNTDFESMPREGRYTGASCVTAACPYNVLT 306
238 PDDTDFACRPHNDGACVPRCPQPLVYNKLTFLQLEPNPHTKYQYGVASCPHNFV-V 296
307 DVGSCCTVCPHNOEVTAEADGTQRCCKSKPCARVCYGLGMEHLREVAVTSANIQEPAG 366
297 DQTSCTVACRPDKMEVD--KNGLKMCEPFGGLCPKACBGTSGS--SRFTVDSNIDGVN 353
367 CKTIFGSLAPLESDGDPASNTAPLQRPOLQVFTLEITGYLYISAMPDLSPLSYFQ 426
354 CTKITGNIDFLITGLNGDPWHNKIPALDEKLVNFTVEITGYLYNIQSPPHMNFVSFS 413
427 NLQVRGELIINGAAS--LTLOGLSWLSGLSLRELSGLALIHNNHLCPVHTYPMQOL 485
414 NLTTGKSLNKGSLIMKMLNVTSLGPNLSLKISGRYISANROLCHHSLSMVKV 473
486 FRNPQALHTA--NRPEDECVGEGLACHQLCARGCMGPRPTQCVNSQFLRGQCEVSEC 544
474 LRGPTEERLDIKHNRPRRDCAVAEGKCDPLCSGGCKMGRPGQCLSCNYSRGCVYHNC 533
545 RVLQGLPREYVNAHCLPCHPECPQNGSVTCFGEADQCVAAAHYKDPPECVACPSGV 604
534 NFLNGEPREFAHAEACFSCHPECCQPMGGATCNCGSDYCAQCAHFRDGPFCVSCPHGV 593
605 KPDLSYMIWKPPEDEGACQPCPINCTHSQVLDLQKGPABQRA----SPLTISIAYVG 660
594 LG--AKGPIYKYPDVQNECRPCHEHNTGCGKGPBLQDCLGQTLVLIGKTHLTMALTVIAG 651
661 ILLVVLGCVGILLIKRRQOKIR--KYTWRRLLQETELVEPLTPSGAMPNOAQMRILKETE 719
652 --LVVIFPMGLGTPLYMGRGRIQNGRAMRRLYERGESIEPLDPS--EKANKYLARI FKE 708
720 LRKAVLCSGAGFYVYKGIWIPDGENVKI PVAIKYLRENTSFPKANKELIDEAYVMAGVGS 779
709 LRKLKVLSSGVGYVHKGVMIPEGESIKI PVICIKYIEKSGROSFOAVTDHMLAGSIDH 768
780 PYSRLSLGICLTSTQVLTOLMPVGCILDHYRENNGRIGSODLIMWCMQIAKMSYLEDV 839
769 AHIVVLGICPSSSIQVLTQVPLGSLLDHVRQHGALGAPOLLNMGVQIAKMYLLEEH 828
840 RLVRHDLAARNVLVSPNHVKITDFGLARLIDIDETEYHADGKVPKIMMALESILRRP 899
829 GMVHRNLAAARNVLKSPGQVADGVADLPRDDKOLLYSBAKPIKIMMALESIHFKY 888
900 THQSDVMSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPORPCTITIDYVMIVK 959
889 THQSDVMSGYVTWELMTFGAEPYAGLRILAEVPLLEKGERLAPQICTIDYVMIVK 948
960 MIDSCRRPRELIVSEFSRMAADPGRFVIVQNEIDGPA---SPLDSTYRSLLDDDDMGD 1016
949 MIDENIRPTFKELANEFTRMADPPRYLVTKRES--GCGIAPGEPHGLTKNKLEVELEP 1007

```

```

1017 LVDAEYLVLPQGFCEPDPAAGA---GMYVHRRHSSTRS-----GGDL 1059
1008 ELDDLDELEAEDMLATTLGASLSPVGLTRPRGCSQLSPSGYPMNQNLGSGSCQ 1067
1060 TLGLEEEERPRPLA-----PSGAGSDVFDGLMGAAKGLS-----LPTHDS 1107
1068 ESAVSGSERCP--PVSLHPNRCGLASESEGHV--TGEAELEKVMCRSRBSRSPR 1125
1108 P-----LQRY--EDTPVPL-----PSETDGYAALPSCPPQEVYNQDPVRPQ--PSP 1152
1126 PRGSAHSGHSHSLTIVTPLSPGLSEEDVNGTV-----MPDTHLKGTSS 1172
1153 REGPLPAARPAAGATLERPKTSPKNGV-----KQVAFAGAVNEPEYLTPOGGAAPQ 1206
1173 REGTLS-----SVGLSVLTGTEBED-------EEYENNRRRRHSPP- 1208
1207 PHPPAPSPAFDNLVYMD-----ODPERGAPSTFKGTPTAENPEYL 1249
1209 PHPPPSLEELGYEYMDVGSLSLGSSTQSCPLHPVPLMPTAGTTPDEDEYEM 1263

```

RESULT 10  
 US-11-113-202-14  
 ; Sequence 14, Application US/11113202  
 ; Publication No. US20050272637A1

```

; GENERAL INFORMATION:
; APPLICANT: Shmiedh, Lara
; TITLE OF INVENTION: RECEPTOR AND METHODS FOR MODULATING SIGNALING BY IGF-1
; FILE REFERENCE: 49321-136
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; SOFTWARE: Patent version 3.3
; LENGTH: 1342
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-113-202-14

```

Query Match 35.8%; Score 2441.5; DB 7; Length 1342;  
 Best Local Similarity 41.0%; Pred. No. 1.3e-147;  
 Matches 539; Conservative 182; Mismatches 457; Indels 137; Gaps 34;

```

10 GLLALLPPGAA--STGYCTGTMKRLPABPETHLDMRLHYGQCQVYQGNLEITYLPT 67
11 GLFLSLARGSEVGNQAVCPGTLNGLSVTGDENQYOTLYKLYERCEVVMGNLEIVLIGH 70
68 NASSLFLDIOEVQGYVLIANQVAVPLQRLRYRGQLEPENVALVLDNGPLNNTT 127
71 NADSLFLQWIREVGYVLIANNEFSTLPNLRVVRGQYVQDGKFAIFVM-----LNNNT 125
128 PVTGASPGLELQRLSLTEILKGGVLIQRPOLCYODTIIMKDI FHNKQALATLIDTN 187
126 ----NSSHALRQLRLTQLTLEILSGGYIEKNDKLCMDTIMRDIVRDRD---AEIVK 178
188 RSRACHPCSPMKSGRCWGESEDCQSLTRTVCAAGC--ARCKGRLPTDCCHGQCAAGCTG 246
179 NGRSCPCHGVCKG--RCWGPGESEDCQTLTKITICAPQCNHCGFGRPNPNCCHDECAAGCGSG 237
247 PKHSPLACLFHNSGICELHCPALVTYNTDFESMPREGRYTGASCVTAACPYNVLT 306
238 PDDTDFACRPHNDGACVPRCPQPLVYNKLTFLQLEPNPHTKYQYGVASCPHNFV-V 296
307 DVGSCCTVCPHNOEVTAEADGTQRCCKSKPCARVCYGLGMEHLREVAVTSANIQEPAG 366
297 DQTSCTVACRPDKMEVD--KNGLKMCEPFGGLCPKACBGTSGS--SRFTVDSNIDGVN 353

```

```

Qy 367 CKKIFGSLAFLPESPDGASNTAPLOPEQLOVETLEETGVLISAMPDLSLDSVQ 426
Db 354 CKKIGLNDLFLTGLNGDPMHKKIPLADPEKLVNFRVRIITGLNIQSPRHNNHNSVYS 413
Qy 427 NLQVIRGRILHNGANS-LTLQIGISWLGIRSLREISGLALIHNTHLCPVHTVPMQDL 485
Db 414 NLTTIGRSLYNGRFSLLIMKINLVNTSLGFRSLKEISASRIYISANROLCYHNSLMTKY 473
Qy 486 FNNPQALLHTA-NREDECVEGGLACHOLCARGHCMGPGPOCVNCSQFLRGQEVENC 544
Db 474 LRGPTBEERLDIKNNRRRRCVAGKVCDDLCSGGCGMGPGQCSNYSRKGVCVTHC 533
Qy 545 RVLQGLPREYVNAHCLPCHPECOPONGSVTCGFPAPDOCVACAHYKDPFVCARCPSCV 604
Db 534 NFLNGEPRFPAHABEFCSCHECOPEGRATOCGSSDVCACAHNRDOPHCSSCPHEV 593
Qy 605 KQDLSYMPKRPDEEGACOPCPINCTHSCVDLDDKXCPAEORA---SPLTSISAVVG 660
Db 594 LG--AKGPIYKYPDVQNECRPCHEMCTQCKGPELDDCLGQTLVLIGKTHLTALVAVIAG 651
Qy 661 ILLVVLGVNPGILLIRROOKIR-KYTMRLLOETLVEPLTPSGAMPQAOIRLIKETE 719
Db 652 --LVVIFMVLGGFTLWRGRRIQNKRAMRYLERGESITRPLDS-BKANKVLARLFEKE 708
Qy 720 LARKVLSGAFGTVYKGIWIPDGBENVKI PVAIKVLRENTSPYANKRIIDEAYVMAVGVS 779
Db 709 LAKVLKSGVGTYNHKGWIPBESIKIPVCIKVIEDSGROSPOAVTDHMLAIGSLH 768
Qy 780 PYVSRLLGICLTSTVOLVQMLPYGCLLDHVRENRRGLSODLLMCMQIAKMSYLEDV 839
Db 769 AHIVRLGLCPSSILQTLVQYLPGLSLDHYVQHRGALPQLLLMNGVLIAGMYVLEBH 828
Qy 840 RLVNHDLAARNVLVSPNHVKTIDPGLABLLIDETRYNADGKVPYIKMALBILRRFP 899
Db 829 GNVHRLAARNVLKSPSQVADLFCVADLLRPDDQQLLYSEAKTPKIMALESIHFGY 888
Qy 900 THQSDVWSYGVYVWELMTFGAKPYDGI PAREIPDLLEKGRILPOPICITIDVYMIWVKCM 959
Db 889 THQSDVWSYGVYVWELMTFGAERYAGLRVLDLEKERRLAPOICITIDVYMWVKCM 948
Qy 960 MIDSECRPRFRLLVSEFSRMAADPORFVVIQNEIDLPA---SPLDSTFYRSLLEDDMGD 1016
Db 949 MIDENIRPTPEKLANEFTMARDPRIYVIKRES--GPGIAPGBEPHGLTNKCLEVELBP 1007
Qy 1017 LVDAEYVLVPOGFCFEPDDAPGA---GGMVHHRHSSSTRS-----CGSL 1059
Db 1008 ELDDLDLDEAEEDNLATTLGSLSLPVGTINRPGSOSILSPSSGYMPOAGNIGGSCQ 1067
Qy 1060 TLGLEPSEBEAPRPLA---PSEGASDVFDGDLGMAKGIQS-----LPTHDS 1107
Db 1068 ESAVSGSSRCPFR-PVSLHMPMRGCLASSESSGCHV-TGSEAELOEKVMSCRSRSRSPR 1125
Qy 1108 P-----LQRY-S-BDPTVPL-----PSETDGYVALPLTCSPOPEYVNOPDVRPQ-PPSP 1152
Db 1126 PRGDSAYHGRSHLPLPTPLSPRGLSEBDVNGYV-----MPDTHLKTPTSS 1172
Qy 1153 REGPRPAARPAATLERPKTSLSPGKGVY-----KDVAFGAVENPEYILPPOGAAPQ 1206
Db 1173 REGTIS-----SVGSSSVLGTSEEBDD-----EYVYNNRRRRHSP- 1208
Qy 1207 PHPPPAFSPAFNLVYWD-----ODPPEKGAPEPTFKGTALENPEYL 1249
Db 1209 PHPPPRSSLEELGYEMDVGSDLSASLGSTGSCPLHPVIMPTAGTTPDEYEM 1263

```

```

; TITLE OF INVENTION: RECEPTOR AND ERBB RECEPTORS
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; PRIOR FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-23

Query Match 27.5%; Score 1874.5; DB 7; Length 419;
Best Local Similarity 89.6%; Pred. No. 2.9e-112;
Matches 353; Conservative 5; Mismatches 29; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLALLPREGASTOYCTGDMKLRLPASPETHLMDLRHLVGGCVVQGNL 60
Db 1 MELAALCRWGLLALLPREGASTOYCTGDMKLRLPASPETHLMDLRHLVGGCVVQGNL 60
Qy 61 ELTYLPTVASLSFLQDIEVQGYVLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120
Db 61 ELTYLPTVASLSFLQDIEVQGYVLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120
Qy 121 DPLNNTPTVPGASPGGLBELQRLSLTEILKGGVLIQRNPQLCYODTIIMKDI FHKNNQIA 180
Db 121 DPLNNTPTVPGASPGGLBELQRLSLTEILKGGVLIQRNPQLCYODTIIMKDI FHKNNQIA 180
Qy 181 LTLIDTNRSRACHPSCPMCKSGRCWGESSESDQSILTRTYCAGGACRCGRLPTDCHBQC 240
Db 181 LTLIDTNRSRACHPSCPMCKSGRCWGESSESDQSILTRTYCAGGACRCGRLPTDCHBQC 240
Qy 241 AAGCTGPKASDCLACLFHNSGICELHCPALVTYNTDFESMNPBEGRTTFGASCVTACP 300
Db 241 AAGCTGPKASDCLACLFHNSGICELHCPALVTYNTDFESMNPBEGRTTFGASCVTACP 300
Qy 301 YNYISTDVSGCTLVCPRLNQEVTAEADGTORCEKSKPCARVCYGLAMEHLREVAVTSAN 360
Db 301 YNYISTDVSGCTLVCPRLNQEVTAEADGTORCEKSKPCARVCYGLAMEHLREVAVTSAN 360
Qy 361 IOEPAGCKKIFGSLAFLPESPDGASNTAPLOP 394
Db 356 LMQPG--PAHPVLSFLRPSWDLVSAFYSLPLAP 387

RESULT 12
US-11-113-202-4
; Sequence 4, Application US/11/113,202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clintcon, Gail M.
; APPLICANT: Shantleh, Lara
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Herstatin; receptor non-binding mutant (Arg to Ile mutation at

```

OTHER INFORMATION: residue 371)  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (371)..(371)  
OTHER INFORMATION: Arg to Ile variant of most common (wild type) sequence  
US-11-113-202-4

Query Match 27.5%; Score 1873.5; DB 7; Length 419;  
Best Local Similarity 89.6%; Pred. No. 3.4e-112;  
Matches 353; Conservative 5; Mismatches 29; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLRPASPEHLDMLRHLVGGCCVVGNTL 60  
DB 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLRPASPEHLDMLRHLVGGCCVVGNTL 60  
QY 61 ETTYPTNASLSFLDIOEVGGVYLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120  
DB 61 ETTYPTNASLSFLDIOEVGGVYLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120  
QY 121 DPLNNTTPTVGTASPGGLRELQRLSTELIKGGVLIQRNPOLCYODTILMKDIFHKNQDLA 180  
DB 121 DPLNNTTPTVGTASPGGLRELQRLSTELIKGGVLIQRNPOLCYODTILMKDIFHKNQDLA 180  
QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGCGARCKGRLPTDCCHQC 240  
DB 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGCGARCKGRLPTDCCHQC 240  
QY 241 AAGCTGPKHSDCLAHFNHSGICEHCPALVTYNTDFPSMNPDEGRYTFGASCVTACP 300  
DB 241 AAGCTGPKHSDCLAHFNHSGICEHCPALVTYNTDFPSMNPDEGRYTFGASCVTACP 300  
QY 301 YNYLSTDVGSCTLVCPPLHNQEVTAEDGTQRCCKSKPCARVYGLGMEHLREVRVTSAN 360  
DB 301 YNYLSTDVGSCTLVCPPLHNQEVTAEDGTQRCCKSKPCARVYGLGMEHLREVRVTSAN 360  
QY 361 IOEPAGCKKIFGSLAFPLPSPFGDDASNTAPLQP 394  
DB 361 IOEPAGCKKIFGSLAFPLPSPFGDDASNTAPLQP 394  
QY 394 LRMQPG--PAHPVLSFLIPSWDLVSAFYSPLAP 387  
DB 394 LRMQPG--PAHPVLSFLIPSWDLVSAFYSPLAP 387

RESULT 13  
US-11-113-202-2  
Sequence 2, Application US/11113202  
Publication No. US20050272637A1  
GENERAL INFORMATION:  
APPLICANT: Clinton, Gail M.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1  
TITLE OF INVENTION: RECEPTOR AND ERBB RECEPTORS  
FILE REFERENCE: 49321-136  
CURRENT APPLICATION NUMBER: US/11/113,202  
CURRENT FILING DATE: 2005-04-22  
PRIOR APPLICATION NUMBER: US 60/590,473  
PRIOR FILING DATE: 2004-07-23  
PRIOR APPLICATION NUMBER: US 60/564,893  
PRIOR FILING DATE: 2004-04-22  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 2  
LENGTH: 419  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (342)..(342)  
OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at this  
OTHER INFORMATION: position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (345)..(345)  
OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at this  
OTHER INFORMATION: position  
FEATURE:

NAME/KEY: MISC FEATURE  
LOCATION: (346)..(346)  
OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this  
OTHER INFORMATION: position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (356)..(356)  
OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this  
OTHER INFORMATION: position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (357)..(357)  
OTHER INFORMATION: Applicants herein disclose Arg and Cys sequence variants at this  
OTHER INFORMATION: position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (358)..(358)  
OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this  
OTHER INFORMATION: position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (361)..(361)  
OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence  
OTHER INFORMATION: variants at this position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (371)..(371)  
OTHER INFORMATION: Applicants herein disclose Arg and Ile sequence variants at this  
OTHER INFORMATION: position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (376)..(376)  
OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this  
OTHER INFORMATION: position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (394)..(394)  
OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this  
OTHER INFORMATION: position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (404)..(404)  
OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this  
OTHER INFORMATION: position  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (413)..(413)  
OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this  
OTHER INFORMATION: position  
US-11-113-202-2

Query Match 27.3%; Score 1860.5; DB 7; Length 419;  
Best Local Similarity 88.8%; Pred. No. 2.3e-111;  
Matches 350; Conservative 3; Mismatches 34; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLRPASPEHLDMLRHLVGGCCVVGNTL 60  
DB 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLRPASPEHLDMLRHLVGGCCVVGNTL 60  
QY 61 ETTYPTNASLSFLDIOEVGGVYLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120  
DB 61 ETTYPTNASLSFLDIOEVGGVYLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120  
QY 121 DPLNNTTPTVGTASPGGLRELQRLSTELIKGGVLIQRNPOLCYODTILMKDIFHKNQDLA 180  
DB 121 DPLNNTTPTVGTASPGGLRELQRLSTELIKGGVLIQRNPOLCYODTILMKDIFHKNQDLA 180  
QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGCGARCKGRLPTDCCHQC 240  
DB 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGCGARCKGRLPTDCCHQC 240  
QY 241 AAGCTGPKHSDCLAHFNHSGICEHCPALVTYNTDFPSMNPDEGRYTFGASCVTACP 300  
DB 241 AAGCTGPKHSDCLAHFNHSGICEHCPALVTYNTDFPSMNPDEGRYTFGASCVTACP 300

Db 241 AAGCTPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACP 300  
QY 301 VNYLSTDVSCITLVNCEPLHNOEVTAEEDTCRCCKSPCARVCYGLGMEHLREVAVTSAN 360  
Db 301 VNYLSTDVSCITLVNCEPLHNOEVTAEEDTCRCCKSPCAR---GXHXXPRDAVVPXX 357  
QY 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLP 394  
Db 358 XQXPRA---HPVLSFLKSMXVSAFYSLPLAP 387

RESULT 14  
US-11-077-386-25

; Sequence 25, Application US/11077386  
; Publication No. US20050272067A1  
; GENERAL INFORMATION:  
; APPLICANT: Macina, Roberto  
; APPLICANT: Turner, Leah R.  
; APPLICANT: Chen, Huel-Mei  
; APPLICANT: Rodriguez, Maria  
; APPLICANT: Liu, Shu-Hui  
; TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Cancer Spec  
; FILE REFERENCE: DEX-0537  
; CURRENT APPLICATION NUMBER: US/11/077,386  
; PRIOR FILING DATE: 2005-03-10  
; PRIOR APPLICATION NUMBER: US 60/566,706  
; PRIOR FILING DATE: 2004-04-30  
; PRIOR APPLICATION NUMBER: US 60/565,144  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/551,911  
; PRIOR FILING DATE: 2004-03-10  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 25  
; LENGTH: 534  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-11-077-386-25

Query Match 15.5%; Score 1056.5; DB 7; Length 534;  
Best Local Similarity 41.1%; Pred. No. 3.4e-60;  
Matches 223; Conservative 76; Mismatches 202; Indels 41; Gaps 13;  
QY 10 GLLALLPPGAA--STVCTGTDMKRLRPASPTLDMRLHYGCGOVVGNLELTYP 67  
Db 11 GLLFSLARGSEVGNQAVCPPTLNGLSYTDABENQYOTLYKLYERCEVVMGMLLETYGR 70  
QY 68 NASLSFLDIOEVOGVVLIANQVQVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTT 127  
Db 71 NADSLFLQIRKREVTGYLVANNEFSTLPPLNLRVVRGTOVDGKFAIFVM-----LNTYT 125  
QY 128 PVTGASPGGLRELQRLSLEILKGGVLIQNRNQLCYODTILMKDIFHKNQALATLIDTN 187  
Db 126 ----NSSHALRQLRLTQLTEILSGGYIEKNKLCMOTIDMRDVRBD---AEIVYKD 178  
QY 188 RSRACHPSPMCKGSRCKWGESSEDCQSLTRITYCAGGC-ARCKGRLPTDCCHQCAAGCTG 246  
Db 179 NGRSCPCHGVCKG-RCWGPGESEDCQTLTKTICAPQCNHCGPMPNQCCHDECGAGGCSG 237  
QY 247 PKHSQCLACIHFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNLYST 306  
Db 228 PDDTCFACRHHNDSGACVPRCPQPLVYNKLTFLQLEPNPHTKYQYGGVCVASCPHNFV-V 296  
QY 307 DVGSCITLVCPILHNOEVTAEEDTCRCCKSPCARVCYGLGMEHLREVAVTSANIQEPAG 366  
Db 297 DQTSQVRACPDPKMEVD-KNGLKMCEPCGGLCPKACBGTGSG--SRFQTVDSNIDGFVN 353  
QY 367 CKKIGSLAFLPESFDGDPASNTAPLPQBPOLQYFETLEITGYLYISANPBDLPULSVQ 426  
Db 354 CTKIIGNLDPLITGLNGDPMKRI PALDPEKLVNFRVIRBITGYLNIQSMPHMHFSVVS 413  
QY 427 NQVIRGRILHNGAVS-LTLQGLGISWLGRLSLRELISGLALIHNTHLCFHTVPMWDL 485

Db 414 NLTTIGSLVYRGFSLLIMKULNVTSLGFSKLKISAGRIYISANROLCYHHSINTKTV 473  
QY 466 FRNPHQALLHRA-NRPDEBCVGE-----GLACHQLCARHCNGPPT 526  
Db 474 LRGPTERRLDIKHNPBRDC-GECKGLGGENRESGRRLKGLFCPRGRSVEGWNQGGG 532  
QY 527 QC 528  
Db 533 GC 534

RESULT 15  
US-11-077-386-26

; Sequence 26, Application US/11077386  
; Publication No. US20050272067A1  
; GENERAL INFORMATION:  
; APPLICANT: Macina, Roberto  
; APPLICANT: Turner, Leah R.  
; APPLICANT: Chen, Huel-Mei  
; APPLICANT: Rodriguez, Maria  
; APPLICANT: Liu, Shu-Hui  
; TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Cancer Spec  
; FILE REFERENCE: DEX-0537  
; CURRENT APPLICATION NUMBER: US/11/077,386  
; PRIOR FILING DATE: 2005-03-10  
; PRIOR APPLICATION NUMBER: US 60/566,706  
; PRIOR FILING DATE: 2004-04-30  
; PRIOR APPLICATION NUMBER: US 60/565,144  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/551,911  
; PRIOR FILING DATE: 2004-03-10  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 26  
; LENGTH: 400  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-11-077-386-26

Query Match 11.7%; Score 798; DB 7; Length 400;  
Best Local Similarity 43.2%; Pred. No. 6.1e-44;  
Matches 163; Conservative 53; Mismatches 141; Indels 20; Gaps 9;  
QY 10 GLLALLPPGAA--STVCTGTDMKRLRPASPTLDMRLHYGCGOVVGNLELTYP 67  
Db 11 GLLFSLARGSEVGNQAVCPPTLNGLSYTDABENQYOTLYKLYERCEVVMGMLLETYGR 70  
QY 68 NASLSFLDIOEVOGVVLIANQVQVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTT 127  
Db 71 NADSLFLQIRKREVTGYLVANNEFSTLPPLNLRVVRGTOVDGKFAIFVM-----LNTYT 125  
QY 128 PVTGASPGGLRELQRLSLEILKGGVLIQNRNQLCYODTILMKDIFHKNQALATLIDTN 187  
Db 126 ----NSSHALRQLRLTQLTEILSGGYIEKNKLCMOTIDMRDVRBD---AEIVYKD 178  
QY 188 RSRACHPSPMCKGSRCKWGESSEDCQSLTRITYCAGGC-ARCKGRLPTDCCHQCAAGCTG 246  
Db 179 NGRSCPCHGVCKG-RCWGPGESEDCQTLTKTICAPQCNHCGPMPNQCCHDECGAGGCSG 237  
QY 247 PKHSQCLACIHFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNLYST 306  
Db 228 PDDTCFACRHHNDSGACVPRCPQPLVYNKLTFLQLEPNPHTKYQYGGVCVASCPHNFV-V 296  
QY 307 DVGSCITLVCPILHNOEVTAEEDTCRCCKSPCARVCYGLGMEHLREVAVTSANIQEPAG 366  
Db 297 DQTSQVRACPDPKMEVD-KNGLKMCEPCGGLCPKACBGTGSG--SRFQTVDSNIDGFVN 353  
QY 367 CKKIGSLAFLPESFDG 383  
Db 354 CTKIIGNLDPLITGLNG 370

Wed Jan 4 10:12:16 2006

us-09-930-125-2.rapbn

Page 12

Search completed: January 3, 2006, 11:28:36  
Job time : 20 secs

---